

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLYSearcher: Al Hansen - EarlySearcher Phone #: 308-4501Searcher Location: Biotech LibDate Searcher Picked Up: 8/24/01Date Completed: 8/27/01

Searcher Prep & Review Time: _____

Clertical Prep Time: 4 minOnline Time: 2 min

Type of Search

NA Sequence (#) 2

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems AB5501

WWW/Internet _____

Other (specify) _____

STIC-Biotech/ChemLib

49545

From: Jiang, Dong
Sent: Thursday, August 23, 2001 4:34 PM
To: STIC-Biotech/ChemLib
Subject: SN09/333,159

6/97
McCarthy, S.A.,
Fraser

Please search SEQ ID NO: 45 and 46

-issued
-commercial

Please send results on paper to Dong Jiang in 10B-01 (mail stop CM1-10C01). Thank you very much.

Dong Jiang (78243)
703-305-1345
U.S. Patent and Trademark Office
Art Unit 1646
dong.jiang@uspto.gov
CM1-10B01
Mail stop: CM1-10C01

Try to find the art only ≥ 54 nt (for 18 aa.)
| 60% var.

RECEIVED
AUG 24 2001
STIC

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:38:56 ; Search time 97.86 Seconds
(without alignments)
2454.893 Million cell updates/sec

Title: US-09-333-159-46

Perfect score: 1269
Sequence: 1 agtgtgaaacctgtcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents.NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appl
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appl
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appl
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appl
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appl
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appl
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appl
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appl
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appl
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appl
11	46.4	3.7	43676	3 US-09-356-952-12	Sequence 12, Appl
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appl
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appl
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 154, App
15	36.6	2.9	3282	2 US-08-276-852-169	Sequence 169, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
17	36.6	2.9	3282	2 US-08-899-575-169	Sequence 169, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
19	36.6	2.9	3282	2 US-08-899-575-169	Sequence 169, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
21	36.6	2.9	3282	5 PCT-US95-08743-169	Sequence 169, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
23	36.6	2.9	13254	1 US-08-276-852-170	Sequence 170, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
25	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
27	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App

28	36.6	2.9	13254	5 PCT-US95-08743-156	Sequence 156, App
29	36.6	2.9	13254	5 PCT-US95-08743-170	Sequence 170, App
30	33.6	2.6	8392	1 US-08-080-255-6	Sequence 6, Appl
31	33.6	2.6	8392	3 US-08-465-713-6	Sequence 6, Appl
32	33.6	2.6	8392	5 PCT-US93-05857-6	Sequence 6, Appl
33	32.6	2.6	7218	1 US-08-232-463-14	Sequence 14, Appl
34	32	2.5	2197	1 US-08-233-005-3	Sequence 3, Appl
35	32	2.5	2197	1 US-08-428-943-3	Sequence 3, Appl
36	32	2.5	2197	3 US-09-016-649-3	Sequence 3, Appl
37	32	2.5	2197	5 PCT-US95-04858-3	Sequence 3, Appl
38	32	2.5	8342	3 US-08-545-860D-63	Sequence 63, Appl
39	32	2.5	8342	5 PCT-US94-04496-63	Sequence 63, Appl
40	31	2.4	3212	3 US-08-673-814-1	Sequence 1, Appl
41	30.4	2.4	1529	1 US-08-336-778-1	Sequence 1, Appl
42	30.2	2.4	4982	3 US-08-699-103B-1	Sequence 1, Appl
43	30.2	2.4	10079	2 US-08-476-866-20	Sequence 20, Appl
44	30	2.4	3639	2 US-08-737-524B-26	Sequence 26, Appl
45	30	2.4	5178	2 US-08-474-169-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-227-108-2
; Sequence 2, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Caps 0;

QY 125 acccagaagcattcatgaattagtgaaatcaatcaacatacaagcctatccctgtgag 184
DB 26 ACCCTGAAGTGACCAATGAATAGTCAGATGATCACCTACTTGGGATACCCAGCTAGG 85
QY 185 aatagaagtcgaactgaagatgggtatattcccttctgtttaacagagattccctcgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGACCGTTATATCTTGGGATCGACAGAAATCCCTTATGGGA 145
QY 245 tagtgaaccttaagaagacagggttccaggcctgtggtgttactcagcagtgccctagt 304
DB 146 GGAATAATTCAGAAATATAGCCGGAGACCTGTGTGATTTTTCACACCGTTGTCTCG 205
QY 305 gaggtgctagcaactgatttccaaactgcccacataagcctgggttccattctggcag 364
DB 206 CATCAGCCACAACTGGATCTCCAACTGCGCAACACAGCCTGGCTTTCATCTGCGCG 265
QY 365 atgctggtttgacgtgtggtgaggggagcagcaggggagcagcctgctcgaaacaca 424
DB 266 ACGCCGGTAGCAGCTGTGCTGGGGAACAGCAGGGGCAACACCTGGCGCAGGAGGAAATC 325
QY 425 agacactctccatagacacagatgagttctgggcttccagtttatgatgagatggttagt 484
DB 326 TGTACTACTCGCCGACTCGCTCGAATCTGGGCTTTTCAGCTTTGACGAGATGCTTAAT 385
QY 485 ttgaccttctcagtgatataaactttatttgacagaaacggccaggaagaaatctatt 544
DB 386 ATGACCTTCGCGCACCAATTTGATTTTTCATCTTGAAGAAACGGGACGACAGATACACT 445
QY 545 atgtcggtctattcacagggcaccacatggcttatttgatgatttccacacatccagagc 604
DB 446 ACGTTGGCCATTTCCAGGGGACCAACCTTGTTCATCGGCTTTTCCACCAATCCCAAGC 505
QY 605 tggctcgaataacaaatgtatttgccttagcaccatagccactgttaagcatgca 664
DB 506 TGCGAAACGGATCAAAACCTTCTATGATTTAGCTCCGTTGCCACCGTGAATACCCG 565
QY 665 aagccccgggacaaattttgtctgccagatgatgataagggatgttttggca 724
DB 566 AAACCTGTGAACAACTCATCTCGTCCCTTCGTTCTTCAAGCTTATATTGGAA 625
QY 725 aaaaagaattctgtatcagacagagatttctcagacactgttatttacccttggcc 784
DB 626 ACAAATATTCTACCCACACCACTTCTTGATCAATTTCTGCCACCGAGGTATGCTCC 685
QY 785 agtgattcttgatcagattgttagtaatatcatgttacttctggtggtgattcaacacca 844
DB 686 GCGAGAGCGTGGATCTCTCTGAGCAACGCCCTGTATTATCATTTTGTGGATTGACACTA 745
QY 845 acaatatgaacatgagccagcaagtgatgatgtgtccacacttctgtggaacatctg 904
DB 746 TGAACCTGAACATGACTCGCTGGATGTGTATCTGTGCATATATCCAGCAGGAAACATCG 805
QY 905 tgcataattctacactgagccagcagtgatgaattctgtgaaactccggcattgact 964
DB 806 TTGAGAACGTGTCTCCATGTCAGGCTGTTAAGTCTGGGAGTTCGAAGTTTGTACT 865
QY 965 gggggagtgagacacaaatctggaaaaatgcaataagcgaactcctgttaaggtacagag 1024
DB 866 GGGGAAGCCAGTTTCAAGACATGATGCATATCATCAGAGCATGCTCCCTACTACAACC 925
QY 1025 tcagagatatgacgggtccctacagacatgtgacagggaggtcaggaactggtttcaaatc 1084
DB 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTGTCTGGCCGACC 985
QY 1085 cagaagacgtgaaatgctgctctctgagtgacacacactcattaccataagaatattc 1144
DB 986 CTCAGCATTTGACCTTTTGGCTTTTCCAGCTCCCAATCTCAATTTACACAGGAAGATTC 1045
QY 1145 ctgaatgggtcagtgagatttcaatctgggttttgatgctcctcaccgtatgtacaatg 1204
DB 1046 CTCCTTAAATCACTTGGACTTATCTGGGCGATGATGCTCCCTCAAGCGGTTTACAATG 1105

QY 1205 aaatcatccatctgtatg 1221
DB 1106 AANTGTTTCCATGATG 1122
RESULT 2
US-09-073-674-2
; Sequence 2, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2

Query Match 35.0%; Score 444.2; DB 2; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 acccagaagcattcatgaattagtgaaatcaatcaacatacaagcctatccctgtgag 184
DB 26 ACCCTGAAGTGACCAATGAATAGTCAGATGATCACCTACTTGGGATACCCAGCTAGG 85
QY 185 aatagaagtcgaactgaagatgggtatattcccttctgtttaacagagattccctcgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGACCGTTATATCTTGGGATCGACAGAAATCCCTTATGGGA 145
QY 245 tagtgaaccttaagaagacagggttccaggcctgtggtgttactcagcagtgccctagt 304
DB 146 GGAATAATTCAGAAATATAGCCGGAGACCTGTGTGATTTTTCACACCGTTGTCTCG 205
QY 305 gaggtgctagcaactgatttccaaactgcccacataagcctgggttccattctggcag 364
DB 206 CATCAGCCACAACTGGATCTCCAACTGCGCAACACAGCCTGGCTTTCATCTGCGCG 265
QY 365 atgctggtttgacgtgtggtgaggggagcagcaggggagcagcctgctcgaaacaca 424


```
Db 266 ACCTGGGTACGACGTGGCTGGGGAACACGAGGGGAACACCTGGCCAGGAGATC 325
QY 425 agacactctccatagaccacgaatgagttctgggtcttcagttatgatgagtgaggt 484
Db 326 TGACTACTCGCCGACCTCCGTCGAATCTCTGGCTTTCAGCTTTCAGGAGATGGTAAT 385
QY 485 ttgacctctcgcagtgataactttatttgcagaaacggccagaaagatctatt 544
Db 386 ATGACCTTCCCGCCACCATGACTTCTTGAAGAAACGGGACGAGCAAGCTACACT 445
QY 545 atgtcggtattccacaggccaccacatgggtttatttgcatttccaccatgccagagc 604
Db 446 AGCTTGGCCATTCCCGAGGACCACTGGTTTCATCGCTTTCACCAATCCCAAGC 505
QY 605 tggctcagaataataatgtatttgggttttagcaccatgacccactgttaagcatgcaa 664
Db 506 TGCGGAACGGATCAAAACCTTCTATGATGATAGCTCCGCTTCCTCTCAAGCTTATTTGGA 625
QY 725 aaaaagaattctgtatcagacacagattcttcagacacattgtatttacccttttgcc 784
Db 626 ACAGAAATATCTACCCACACCACTCTTTTGATCAATTTCTGCCACCGATGCTGCC 585
QY 785 aggtgattcttgcagatgtttagtaataatcttcttcttcttcttcttcttcttcttctt 844
Db 686 GCGAGACGCTGATCTCTCTCAGCAACGCGCTCTTTATCATCTTTGATTTGACACTA 745
QY 845 acaatatgaacatgacgagcaagtgatgtatgtctgcccacacttcttctggaacatctg 904
Db 746 TGAACCTTGAACATGAGTCTGGATGTGTCTGTCTACATTAATCCAGCAGGAACATCGG 805
QY 905 tgcataaattattctacactgagccagcgagtgaaattctggtgaactcgggactttagct 964
Db 806 TTCAGAACGCTGCTCCACTGGTCCAGGCTGTTAAGCTCTGGGAAGTTCCAAGCTTTGACT 865
QY 965 gggggagtgagaccacaaatctggaataatgcaatcagccaaactctctgttaaggtacag 1024
Db 866 GGGGAAGCCAGTTTCAGAACATGATGACATATCATCATCAGACGATGCTCTCTACTACAAC 925
QY 1025 tcagagatatgaggtccctacagcaatgtgacagagaggtcaggaactgttcaaatc 1084
Db 926 TGACAGACATGATGTCGAATTCGCAAGTGTGGAACGGTGGCAACGACTGTGCGCCGACC 985
QY 1085 cagaacagctgaaatgctgctctctgagtgaccacacacacacacacacacacacacac 1144
Db 986 CTCACGATGTTGACCTTTTCTTCCAGCTCCCAATCTCAATTTACCAAGGAAGATC 1045
QY 1145 ctgaatgggctcagctgggatttctcctggttgggttgggtgctcctcaccgtatgcaatg 1204
Db 1046 CTCCTTACAATCATTGGACTTTATCTGGCCCATGATGCTCCCTCAAGCGGTTTACAATG 1105
QY 1205 aaatcatcatctgatg 1221
Db 1106 AAATTTGTTCCATGATG 1122
```

RESULT 3

US-08-227-108-4

; Sequence 4, Application US/08227108

; Patent No. 5807726

; GENERAL INFORMATION:

; APPLICANT: Blanchard, Claire

; APPLICANT: Benicourt, Claude

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSES: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

```
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-08-227-108-4
```

Query Match 35.0%; Score 444.2; DB 1; Length 1140;

Best Local Similarity 62.8%; Pred. No. 5.6e-139;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

```
QY 125 acccagaagcattctatgaataattagtgaaatcatcacaatcatcaaggtctccctgtgag 184
Db 29 ACCCTGGAAGTGACCATGAATATAGTCAGATGATCACCCTACTGGGGATACCCAGCTGAGG 88
QY 185 aatagaaagtcgaactgaagatgggtatctcttctgttaacaggattctctcgagcc 244
Db 89 AATATGAAAGTTGTGACCGGAAGACGGTTATATCTCTGGGATCAGACAAATCTCTATGGGA 148
QY 245 tagtgcaacctaaagaagacaggttccagcctgtggtgttactgcagcagtggtcctagt 304
Db 149 GGAATAATTCAGAGAAATATAGCCCGGAGACCTGTTGCATTTTTCACACAGGTTTCTCTCG 208
QY 305 gagjtgctagcaactggatttccaaacctgtcccaacaatagcctgggtcttcattctctg 364
Db 209 CATAGCCACAACTGGATCTCCAACTGCCCCAACACAGCCTGGGCTTTCATCTCTGGCCG 268
QY 365 atgtcggttttgacgtgtggtgaggaagacagcaggggaaacgcctgtgtctcgaacaca 424
Db 269 ACGCCGGGTACGACGTGTGGCTGGGGAACAGCAGGGGGCAACACCTGGGCCAGGAGAAATC 328
QY 425 agacactctccatagaccaagatgagttctgggtcttcagttatgatgagtggtcaggt 484
Db 329 TGTACTACTTCGCCCGACACTCCGTCGGAATCTCTGGGCTTTTCAGCTTTTGACGAGATGGCTAAAT 388
QY 485 ttgacctctcctcagtgatataaactttatttgcagaaacggccaggaagaaatctatt 544
Db 389 ATGACCTTCCCGCCACCATGACTTCTATCTTGAAGAAACGGGACGAGCAAGCTACACT 448
QY 545 atgtcggtattccacaggccaccacatggtttatttgcatttccaccatgcccagcagc 604
Db 449 ACGTTGGCCATTTCCAGGGCACCACCATTTGGTTTTCATCGCTTTTTCACCAATTCACCAAGC 508
QY 605 tggctcagaataataatgtatttgggttttagcaccatgacccactgttaagcatgcaac 664
Db 509 TGGCGAAACGGATCAAAACCTTCTATGATGATAGCTCCGCTTCCCGGTAAGATACACCG 568
```


Db	809	TTCAGAACGTGTCACACGTGGTCCAGAGCGTTAAAGTCTGGGAAGTTCACAGCTTTTGACT	868
Qy	965	gggggagtgagaccacaaaactctggaaaaatgcaatcatgaccaactcctctgaagtacagag	1024
Db	869	GGGGAGGCCAGTTCAGAACATGATGACATCATCAGACGTCCTCCCTACTACAACC	928
Qy	1025	tcagagatatgacggtccctacagcaatgagacaggaggtcgaggactggcttctcaaatc	1084
Db	929	TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTCTGGCCGACC	988
Qy	1085	cagaagagtgaaaatgctgctctctgaggtgacaaacctcatctaccataaagaatttc	1144
Db	989	CTCAGCATGTTGACCTTTGCTTTCCAAAGTCCCAATCTCATTTACACAGGAAGATTC	1048
Qy	1145	cigaatggggtcacgtgattctcatctggggttggatgctccctcaccgctatgtacaatg	1204
Db	1049	CTCCCTTACAACTACCTTGACATTATCTGGGGCCATGGATGCCCTTCAGCGGTTACATG	1108
Qy	1205	aaatcatcatctgatg	1221
Db	1109	AAATTGTTTCCATGATG	1125

RESULTS

```

US-08-227-108-6
: Sequence 6, Application US/08227108
: Patent No. 5807726
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/227,108
: FILING DATE: 03-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Fanucci, Allan A.
: REGISTRATION NUMBER: 30,256
: REFERENCE/DOCKET NUMBER: 7620-033
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090
: TELEFAX: 212 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1146 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: US-08-227-108-6

Query Match          35.0%; Score 444.2; DB 1; Length 1146;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps
QY 125 acccagaagcattcatgaattattgaattgaatcatcccaacatcaagactatccgtatgcag 184

```


TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-09-073-674-1

Query Match 35.0%; Score 444.2; DB 2; Length 1528;
Best Local Similarity 62.8%; Pred. No. 6.8e-139; Mismatches 0; Indels 0; Gaps 0;
Matches 689; Conservative 0;

QY 125 accagaagcattcatgaattatttagtgaatcatcatcaacatcaatcaagcctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATCAATATAAGTCAGATGATACCTACTGGGATACCCAGCTGAGG 85

QY 185 aatatgaagtgcgaactgaagtggtatataatcccttctgttaacagagattccctgagcc 244
DB 86 AATATGAAGTTGTGACCGAAGCGGTATATCCTTGGGATCGACAGAAATTCCTTATGGGA 145

QY 245 tagtgaacctgaagagacaggttccagcctgtgtgttactgcagcagtgccctagt 304
DB 146 GGAATAATTCAGAGAAATATAGCGCGGAGACCTGTTCATTTTGCACACAGGTTTCTCG 205

QY 305 gagggttagcactgatttccaaactgcccacaaatagcctgggcttcaattctggcag 364
DB 206 CATCAGCCACAACTGGATCTCAACTGCCAACACAGCGCTGCTTCATCCTGGCG 265

QY 365 atgctggttttgacgtgtgtgagtgagagcaggggaaacgcctggtctgaaacaca 424
DB 266 ACGCCGGGTACGACGTGTGCTGGGACAGCAGGAGGCAACACTGGCGCAGGAGAAATC 325

QY 425 agacactctccatagaccagaatgttctgtgggtttcagttatgatgagtggttagt 484
DB 326 TGACTACTGCGCCGACCTCGTGAATTTGGGCTTTTACGCTTTGACGAGATGGCTAAAT 385

QY 485 ttgacctctctgagtgataaaattttatttgcagaaacggccaggaagaaatcatt 544
DB 386 ATGACCTTCCCGCCACCATGTACTTCTTGAGAAACGGGACAGACAGCTACAT 445

QY 545 atgctggtatttcacagggccacacacatgggtttatttgcattttccaccatgcagagc 604
DB 446 ACCTTGGCCATTTCCAGGGCACCACCATTTGTTTCAATCGCTTTTCCACCAATCCCAAGC 505

QY 605 tggctcagaaatcaaaatgtatttgccttttagcaccatagccactgttaagcatgcaa 664
DB 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGCTCCCTGCGCCAGGTGAAGTACACCG 565

QY 665 aaagcccccgggaccaaattttgttgcagagatgatgatcaaggagattgttggca 724
DB 566 AAACCTGTTAAACAACTCATGCTCGTCCCTTCGTTCTCTCAAGCTTATTTGGAA 625

QY 725 aaaaaaattctgtatcagaccagatttctcagacaaattgttattaccctttgtggcc 784
DB 626 ACAAATATTTCTACCCACACCACTTCTTTGATCAATTTCTCGCCACCGAGGTATGCTCC 685

QY 785 aggtgattctgacagattgttagtaataatcatgttaattcttctggtgattcaacacca 844
DB 686 CGGAGACGGTGTATCTCTGTCAGCAACGCCCTGTTTATCATTTTGTGGATTGACACTA 745

QY 845 acaatatgaacatgagccagcaagaagtgtatgtctccacacactcttctggaacatctg 904
DB 746 TGAACCTGAACATGATCGCTTGGATGTGTATCTGTCTACATATCCAGCAGGAACATCGG 805

QY 905 tgcataattctacactggagccagcagtggaattctggtgaactcccgccatttgaact 964
DB 806 TTCAGAACGTGCTCCACTGTGCTCCAGGCTGTTAAGTCTGGGAAGTTCCCAAGCTTTTGACT 865

QY 965 gggggagtgagacacaaatctggaataatgcaatcagcacaactcctgttaaggtacagag 1024
DB 866 GGGGAAGCCCGATTCAGACATGATGCATATCATCAGAGCATGCTCCTCCTACTACACC 925

QY 1025 tcagagatatgacgttccctacagcagatgtggacagaggtcagggactggcttcaaatc 1084
DB 926 TGACAGACATGCATGTGCCAATCGCAGTGTGGNACGGTGGCAACGACTTGTCTGGCGGACC 985

QY 1085 cagaagacgtgaaaatgctgctctctgaggtgaccacacctcctaccataagaatattc 1144
DB 986 CTCAGGATGTTGACCTTTTGTCTTCCCAAGCTCCCAATCTCATTTACCACAGGAATTC 1045

QY 1145 ctgaatgggctcactggtgatttcactctctggtggttggatgctctcaccgtatgacaatg 1204
DB 1046 CTTCTTACAATCACTTGGACTTTATCTGGCCATGATGCCCTCAAGCGGTTTACAATG 1105

QY 1205 aaatcatccatctgatg 1221
DB 1106 AAATGCTTCCATGATG 1122

RESULT 9
US-08-751-782-2
Sequence 2, Application US/08751782
Patent No. 5821352
GENERAL INFORMATION:
APPLICANT: Heintz, Nathaniel
APPLICANT: Gubbay, Johnathan
APPLICANT: Skinner, Michael
TITLE OF INVENTION: A cDNA Library Prepared during
Regression of Rat Prostate and Methods of Use Thereof
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,782
FILING DATE: 18-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-190
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 178 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
IMMEDIATE SOURCE:
CLONE: 10.2
US-08-751-782-2

Query Match 3.8%; Score 48.2; DB 1; Length 178;
Best Local Similarity 55.8%; Pred. No. 1.4e-06;
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 416 gaaacacacagacactctccatcagaccagagtggtctgggcttctcagttatgagaga 475
DB 8 GGAACATGTGGCGCTTAGACCCAGATTCCTAAAGAAATTTGGGATTTAGTTTAAATGAAC 67

QY 476 tggctaggtttacccttctcctgcagtgataaaacttttttgcagaaaaacggccagaaa 535
DB 68 AAATAGATACGACCTCCCGCCATCATTTATTTCATCTGTAATGAACAAGCAACAC 127

RESULT 13
 US-09-163-669-1
 ; Sequence 1, Application US/09163669
 ; Patent No. 6111076
 ; GENERAL INFORMATION:
 ; APPLICANT: FUKUSUMI, SHOJI
 ; APPLICANT: HINUMA, SHOJI
 ; APPLICANT: FUJII, RYO
 ; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTOR (HIBCD07)
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Ragner & Prestlia
 ; STREET: P.O. Box 980
 ; CITY: Valley Forge

GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TO HUMAN IMMUNODEFICIENCY VIRUS
DATE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:

Query Match	2.9%	Score 36.6;	DB 1;	Length 3282;
Best Local Similarity	51.5%;	Pred. No. 0.074;		
Matches 84;	Conservative	0;	Mismatches 79;	Indels 0;
				Gaps 0;

[illegible]

RESULT 15
us-08-276-852-169/c
Sequence 169, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:44:11 ; Search time 4395.77 Seconds
(without alignments)
4465.332 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgtcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_ba3:
4: gb_in1:
5: gb_in2:
6: gb_in3:
7: gb_om:
8: gb_ov:
9: gb_pat1:
10: gb_pat2:
11: gb_ph:
12: gb_pl1:
13: gb_pl2:
14: gb_pl3:
15: gb_pl4:
16: em_ba1:
17: em_ba2:
18: em_fun:
19: em_htgo_hum:
20: em_htgo_inv:
21: em_htgo_rod:
22: em_htg_hum1:
23: em_htg_hum2:
24: em_htg_hum3:
25: em_htg_hum4:
26: em_htg_hum5:
27: em_htg_hum6:
28: em_htg_hum7:
29: em_htg_hum8:
30: em_htg_inv1:
31: em_htg_inv2:
32: em_htg_other:
33: em_htg_rod:
34: em_hum1:
35: em_hum2:
36: em_hum3:
37: em_hum4:
38: em_hum5:
39: em_hum6:
40: em_hum7:
41: em_in:
42: em_om:
43: em_or:

44: em_ov:
45: em_pat:
46: em_ph:
47: em_pl:
48: em_ro:
49: em_sts:
50: em_sy:
51: em_un:
52: em_v1:
53: gb_sts1:
54: gb_sts2:
55: gb_sts3:
56: gb_sy:
57: gb_un:
58: gb_v11:
59: gb_v12:
60: gb_htg1:
61: gb_htg2:
62: gb_htg3:
63: gb_htg4:
64: gb_htg5:
65: gb_htg6:
66: gb_htg7:
67: gb_htg8:
68: gb_htg9:
69: gb_htg10:
70: gb_htg11:
71: gb_htg12:
72: gb_htg13:
73: gb_htg14:
74: gb_htg15:
75: gb_htg16:
76: gb_htg17:
77: gb_htg18:
78: gb_htg19:
79: gb_htg20:
80: gb_htg21:
81: gb_htg22:
82: gb_htg23:
83: gb_htg24:
84: gb_htg25:
85: gb_pr1:
86: gb_pr2:
87: gb_pr3:
88: gb_pr4:
89: gb_pr5:
90: gb_pr6:
91: gb_pr7:
92: gb_pr8:
93: gb_pr9:
94: gb_rol:
95: gb_rod:
96: gb_in4:
97: gb_prio:
98: em_ba3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	527.8	41.6	2481	93	HSYACLY
2	527.8	41.6	2481	97	HSU08464
3	527.8	41.6	2626	93	HSIAL
4	526.2	41.5	2493	97	HUMLIPCHL
5	489.2	38.6	1378	9	A26689
6	489.2	38.6	1378	9	A26690
7	470	37.0	1365	93	HSGLR
8	470	37.0	1367	9	A01046

231690 H.sapiens (U08464 Human lysos
X76488 H.sapiens m
M74775 Human lysos
A26689 Precursor o
A26690 Precursor o
X05997 Human mRNA
A01046 H.sapiens m

QY 513 ttgcagaaacagggccaggaagaaatctattatgtcgctattccagagggcaccacat 572
Db 510 TCTGAATAAACTGGCCAGAAAGAGTATTATGTGGTCAATCTCAAGGCCACTAT 569
QY 573 gggcttattgctatttccaccatgcagagctgcagaaatcaaaatgtatttgc 632
Db 570 AGGTTTATAGCATTTTCAGATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTGCG 629
QY 633 tttagcaccatagccactgttaagcatgcacaaagcccgagcacaatttttgttgc 692
Db 630 CCTGGGCTCTGGCTCGCTCCCTCTGTACTAGCCCTATGGCCAAATAGGACGATT 689
QY 693 gccagatatgatcatcaaggatttggcacaagaaatcttctatcagaccagatt 752
Db 690 ACCAGATCATCTCATTAAAGACTTATTGGAGACAAGAAATTTCTCCCGAGAGTCGGT 749
QY 753 tctcagaacactgttatttacccttggcaggtgattcttctgacagatttggtagtaa 812
Db 750 TTTGAAGTGGCTGGTACCCACGCTTTGCACATCATGTCACTGAAGGAGCTCTGTGAAA 809
QY 813 tateatgttactctctgggtggtatcacaacacaaatatagaacatgagccagagtg 872
Db 810 TCCTGTTTCTCTGCTGTGGATTAAAGAGAGAAATTTAAATATGCTAGAGTGGATG 869
QY 873 atagctgcccacactcttctggaacatctgtgcaaaatattctacactggagccaggc 932
Db 870 ATATACAACACATCTCCTCTGGAACTTCTGTGCAAAACATCTTACACTGGAGCCAGGC 929
QY 933 agtgaattctgtgaaactccggcatttgaactggggagtgagacacaaatctggaaa 992
Db 930 TGTAAATTCCAAAGTTTCAAGCCTTGAAGCTGGGAGACAGTCCCAAGATTTATTTC 989
QY 993 atgcaatcagcaccactctgtaaggtacagagtcagagatgacaggtccctacagcaat 1052
Db 990 TTACAACACAGATTATCTCCACATACATATGTAAGGACATGCTTGTGCGGACTCGAT 1049
QY 1053 gtgacagaggtcagagctggttccaaatccagagacgtgaaatgctgctctga 1112
Db 1050 CTGAGCGGGGTCAGACTGGCTGCAGATGCTACGAGCTCAATATCTTACTGACTCA 1109
QY 1113 ggtgaccaactctatcaccataagaattctctgaatgggtccagtgattctatctg 1172
Db 1110 GATCACCACATGTTGCTTCCATGAGACATTCGGGAATGGGACATCTTGACTTCATTG 1169
QY 1173 ggttggatgctctcaccgtatgtacaatgaatcattccatctgagcagcagagga 1232
Db 1170 GGGCTGGATGCCCTTGGAGGCTTTATATAAAATATTAAATCTAATGAGGAATATCA 1229
QY 1233 gaccacact 1241
Db 1230 GTGAAGCT 1238

RESULT 2
HSU08464
LOCUS HSU08464 2481 bp mRNA PRI 23-JUN-1994
DEFINITION Human lysosomal acid lipase mRNA, complete cds.
ACCESSION U08464
VERSION U08464.1 GI:505052
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2481)
AUTHORS Du,H. and Gregory,G.A.
TITLE Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2481)
AUTHORS Du,H.
TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Hong Du, Division of Human Genetics,

Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH 45229-3039, USA
FEATURES
Location/Qualifiers
1..2481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="clone HLAL"
/clone_lib="HepG2 cDNA library/Lambda 2AP"
/cell_line="HepG2"
/cell_type="hepatoma"
/tissue_type="liver"
5'UTR
1..33
34..96
sig_peptide
CDS
34..1233
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"
/translation="MKMRPLGLVCLVLTWHSEGRGKLTAVDPETNMNVSEIISW
GFPSEELVETEDGYILCUNRIPHRKKNHSDGPKPVVFLQHGLLADSSNWTNLANS
SLGFLADAGFDVWNGNSRGTWKRHKLTLSQDEFWAFSYDEMAKYDLPASINFL
NKTGOEVVYVGHSGQTTIGTAFSQIPELAKIRKMFALGPVAVSFCTSPMAKLR
LPDHLIKDLFGKDFELPQSAFLKWLGTHTVTKILKELCNLCFLICGNERNLNMR
VDVYTHSPACTSVQNMHLHWSQAVKFKQAFDWSGSSAKNYEHYNSQYPTYNVKDML
VPTAVWSGHDWLADYIDVNIILTLQINDVHFHSEIPWEHDLDFIWGLDAPWRLINKII
NLMRYQ"
variation
79
/replace="c"
mat_peptide
97..1230
/product="lysosomal acid lipase"
variation
100
/replace="g"
3'UTR
1234..2480
polya_signal
2459..2464
polya_site
2480
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN
Query Match 41.6%; Score 527.8; DB 97; Length 2481;
Best Local Similarity 65.4%; Pred. No. 4.5e-146;
Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1;
QY 39 cagaatggaaatgtgcttctgtattctgtggtggtggtatattgtccagagaaatggaattc 98
Db 30 CAGAATGAAAATCGGCTTCTTGGGTGGTGGTGTGTGTGTCTCTGGACCTGCATTC 89
QY 99 ag-----tacatatgccaaactaaagctgtggacccagagaagcattcatgaattagta 152
Db 90 TGAGGGGTCTAGAGGAAAACCTGACAGCTGTGGATCTCTGAACAAACATGAATGTGAGTGA 149
QY 153 aatcatccaacatcaaggctatccctgtgaggaatatgaagtcgcaactgaagatgggta 212
Db 150 AATTATCTTACTTGGGATTCCTCTAGTAGGAATACCTAGTGAGACACAGATGGATA 209
QY 213 tatectttctgttaacagattcctctgagggccttagtgcacacctaagaagacaggttccag 272
Db 210 TATTCTGTGCTTAACCGAATTCCTCATGGAGGAAGAACCATCTTGACAAAGGTCCCAA 269
QY 273 gectgtggttactcagcagcattgctgaggtgtgaggtgtcagcaactggatttccacact 332
Db 270 ACCAGTTGCTTCTCTCAACATCGCTTGTCTGCGAGATTCCTAGTAACCTGGGTCAAAACCT 329
QY 333 gccacaataagcctgggtctcattcggcagatgctggttttgacctggtgagtgaggaa 392
Db 330 TGCCACAGCAGCCTGGGCTTCATCTTGTCTGATGTGGTGTTCACGTGTGGATGGGCAA 389
QY 393 cagcaggggaaacgcctggtcctcgaaaaacacagacactctccatagaccaagatgatt 452
Db 390 CAGCAGGAAATACCTGGTCTCGGAACATAGACACTCTCAGTTCTCAGATGAATT 449
QY 453 ctggccttccagtattgatgagatgaggttttgaccttgcagtgatgataaacttat 512
Db 453 ctggccttccagtattgatgagatgaggttttgaccttgcagtgatgataaacttat 512

Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH 45229-3039, USA
FEATURES
Location/Qualifiers
1..2481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="clone HLAL"
/clone_lib="HepG2 cDNA library/Lambda 2AP"
/cell_line="HepG2"
/cell_type="hepatoma"
/tissue_type="liver"

AUTHORS Anels, D.
TITLE Direct Submission
JOURNAL Submitted (29-NOV-1993) D. Anels, Medical Department, University
FEATURES Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG
source Location/Qualifiers
1. .2626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/tissue_type="liver"
146. .1345
/gene="LAL/LIPA"
146. .1345
/gene="LAL/LIPA"
/EC_number="3.1.1.13"
/notes="sterol esterase"
/codon_start=1
/evidence=experimental
/product="lysosomal acid lipase"
/protein_id="CAA54026.1"
/db_xref="GI:434306"
/db_xref="SWISS-PROT:P38571"
/translation="MKRFLGLVVLVLTNLHSEGGSGKLTAVDPETNMNVSEIISYWG
GFSEEEVLVETEDGYILCLNRIPHGRKHNKSDGPKPVFLQHGLLADSNNVTNLAIS
GFEIILADAGDFVYMGNSYTRSRKHTLSVSQDEFWAFSYDEMAKYDILPASINFLIT
NKYGEQVYVGHSGQGTIGFIASFQIPELAKIKMFALGCPVAFASVAFCTSPMAKLGSR
LPDHLIKDLFLKQEFEPQSAFLKWLGHVCTHVLKELCGNLCLFLCGFERNLNLRMSR
LDVYTHSPAGTSVONMLHWSQAVKFOKFOAFDAGSSAKNYFHYNOSYPTTYNVKDM
VPTAVSGGHGHWLADYVDVWILLTQITNLVHESIPWEVHLDFIWLGLDAPWRLYNKIKI
NLMRKRYQ"
sig_peptide 146. .226
/gene="LAL/LIPA"
/evidence=experimental
227. 373
/gene="LAL/LIPA"
/notes="propeptide"
/evidence=experimental
374. .1342
/gene="LAL/LIPA"
/evidence=experimental
/product="lysosomal acid lipase"
BASE COUNT 741 a 527 c 547 g 811 t
ORIGIN
Query Match 41.6%; Score 527.8; DB 93; Length 2626;
Best Local Similarity 65.4%; Pred. No. 4.6e-146;
Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1
Qy 39 cagaatgaaatggtgctctgattctctgtggtggtggtatgtgtccagagaaatgtgaattc 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 CAGATGAANAATGCGGTCTCTGGGGTGTGGTGTCTGTGTGGTCTCTCTGGACCTGCATTC 201
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 99 ag-----tacatatgccaaactaaagctgtggaccagagaagcattcatgaattagtga 152
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 TGAGGGGTCTGGAGGAAACTCACAGCTGTGGATCCTGAAACAAACATGAATGTGAGTGA 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 153 aatctccaaatcaagctatccctgtgaggaaatataaagtcgcaactgaagatgggta 212
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AATTATCTCTTACTGGGGATTCCCTAGTAGGGAATACCTAGTTGAGACAGAGATGGATA 321
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 213 tctctttctgtaacagattcttcgaggtcctagtcgaacctagaagacaggttccag 272
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 TATTCTGTGCCCTAACCGAATTCCTCATTGGGAGGAAGAACCAATCTGACAAAGGTCCCA 381
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 273 gcctgtggttactgcagcatggcctcctagttgaggtgctagcaactggatttccaaact 332
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 ACCAATGTGCTTCCTGCAACATGGCTTGCTGCGAGATTCTACTAACTGGGTACAAACCT 441
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 333 gcccaacaatagcctgggtcttcattctctgagagatcgtggttttgacgtgagtgagga 392
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 TGCCACAGCAGCCCTGGGCTTCATTCTCTGCTGATGCTGTTTGTACGCTGTGATGGCCAA 501
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 217 TATTCTGCTTAAACCAATTCCTCATGAGGAGGAAGAACCAATTCCTGACAAAGGTCCTCAA 276
 Qy 273 gcctgtgttactgacagcatggccttagtgagggtgctagcaaatggtatttcccaacct 332
 Db 277 ACCAGTGTCTCTCTGCAACATGGCTGCTGGCAGATTCCTAGTAACCTGGGTCAAAACCT 336
 Qy 333 gcccaacaatagcctggcttctctgagcatgctggttttgcagctggtgagtgaggaa 392
 Db 337 TGCCACAGCAGCGCTGGCTTCATCTCTGCTGATGCTGGTTTTCAGCTGGGATGGCAA 396
 Qy 393 cagcaggggaaacgctggcttcgaaacacagacactctccatagacacaaagatgatt 452
 Db 397 CAGCAGAGGAATACCTGGTTCGGAACATATAGACACTCTCAGTTCTCAGATGAAT 456
 Qy 453 ctgggttccagtattgatgagatggctaggtttgaccttctcgcagtgataaaactttat 512
 Db 457 CTGGCTTTCAGTATGATGAGATGGCAAAATATGACCTACCACTTCCATTAACCTCAT 516
 Qy 513 ttgcagaaaaacggcaggaaagatctattatgctggctattacagaggccaccacct 572
 Db 517 TCTGAATAAACTGGCCCAAGCAAGTATATATGTTGGTCTATCTCAAGGCCACCACTAT 576
 Qy 573 gggctttatgctatttccaccatgcccagagctggctcagaaaaatcaaatgtatttgc 632
 Db 577 AGGTTTATAGCATTTTCAGATCCCTGAGCTGGCTTAAAGGNTTAAATGTTTTTTCG 636
 Qy 633 tttagcaccatagccactgttaagcatgcaaaaagccccgggaccaaatttttggct 692
 Db 637 CCTGGGTCCTGTGGCTTCCGCTCTCTCTAGTACGCTATGGCCAAATTAGGACGATT 696
 Qy 693 gccagatgatgatcagaagatttggcgaagaaatcttctgtatcagaccagatt 752
 Db 697 ACCAGATCTCTNTAAGACTTATTGGAGACAAAGAAATTTCTTCCCGAGAGTGGGTT 756
 Qy 753 tctcagaaactgttatttaccctttgtggcgggtgattcttgatcagattttagtaa 812
 Db 757 TTCAAGTGGCTGGGTACCCACAGTTTGCACATGTCATCTGAAGGAGCTCTGTGAAA 816
 Qy 813 tateatgttacttctgggtgatttcaacacacaaatatagaacatgagccgagcaagt 872
 Db 817 TCTCTGTTTCTCTGTGTGGATTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATGT 876
 Qy 873 atatgctgccacactcttctggaacatctgtgcaaaatattctacactggagccagc 932
 Db 877 ATATACAAACATTTCTCTCTGGAATCTGTGCAAAACATGTTACACTGGAGCCAGC 936
 Qy 933 agtgaattctggtaactccggcatttgcattgagtgagggtgagagcaaaatctggaaa 992
 Db 937 TGTTAAATTCAAAAGTTTCAAGCTTTGACTGGGAAGCAGTCCCAAGATATTATTCA 996
 Qy 993 atgcaatcagccaaactctgtaagggtacagagtcagagatatgacgggtccctacagcaat 1052
 Db 997 TTACAACAGAGTATCTCTCCACATACATATGTGAAGGACATGCTTGTGCGACTGCAGT 1056
 Qy 1053 gtgacagagggtcagactggtttcaaatccagaagacgtgaaatgctgctctga 1112
 Db 1057 CTGAGCGGGGGTCCACACTGGGCTTGCAGATGTCTAGACGTCAATATCTTACTGACTCA 1116
 Qy 1113 ggtgaccaactcctatcaccataagaatatctcgaatgggtcagctgatttcatctg 1172
 Db 1117 GATCAACCACTTGGTGTTCATGAGACATTCGGGAATGGGAGCATCTTCACTTCATTG 1176
 Qy 1173 gggtttggatgctcctcaccgtagtgataatgaatccatcctgatgagcagagga 1232
 Db 1177 GGGCTGGATGCCCTTGGAGGCTTTATATAAAATATTAAATATTAAATGAGGAATATCA 1236
 Qy 1233 gaccaacct 1241
 Db 1237 GTGAAGACT 1245

RESULT 5
 A26689

LOCUS A26689 1378 bp DNA PAT 05-APR-1995
 DEFINITION Precursor of rabbit gastric lipase coding sequence.
 ACCESSION A26689
 VERSION A26689.1 GI:905029
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 1378).
 AUTHORS Benicourt, C., Blanchard, C. and Junlen, J.L.
 TITLE Recombinant gastric lipase from rabbit and pharmaceutical compositions.
 JOURNAL Patent: EP 0542629-A 9 19-MAY-1993;
 INSTITUT DE RECHERCHE JOUVEINAL
 FEATURES
 source
 1. .1378
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 397 a 287 c 276 g 418 t
 ORIGIN
 Query Match 38.6%; Score 489.2; DB 9; Length 1378;
 Best Local Similarity 65.0%; Pred. No. 1.4e-334;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 QY 125 acccaagaacatcatcgaatatttagtgaatcatcccaacatcaaggctatccctgtgagg 184
 Db 136 ACCCIGAAGTGAATATGAATATTAGTCAGATGATTCTTACTGGGATACCCAGTGAAA 195
 QY 185 aatataagtcgaactgaagatgggtatatctcttctgtaacagagattcctcgagcc 244
 Db 196 AATATGAAGTGTGACTGAAGATGGTTATATCTCTAGTTTACAGAAATTCCTTATGGA 255
 QY 245 tagtgaacctaaagaacagagttccagcctgtggtgttactcagacagcctagttg 304
 Db 256 AGAAATTCAGGACAGAGGCGCCAGACCCGCTGATTTTTCAGCATGGTTGCTTG 315
 QY 305 gagggtctagcaactggatttccaaacctgcccaacatagcctgggttcttattctgagc 364
 Db 316 CATCAGCCTCAAACTGGATCTCCAACTGCCCAACACAGCGCTGGCTTTATTCTGCAG 375
 QY 365 atgctgggtttgacgtgtgagtgagggaacgaggggaaacgctgtctcgaacacaa 424
 Db 376 ATGCTGGTTATGTGTGTGGCTGGGGAACAGCAGAGAAATACCTGTGTCCAGAGAACT 435
 QY 425 agacactctcagacacaaagatgagttctgggctttcagttatgatgagatgctaggt 484
 Db 436 TATATATTCCAGCATCTCGTTGAATTCGGGCTTTTAGCTTTGATGAATGGCTAAAT 495
 QY 485 ttgaaccttctcagtgatgaataactttattttgcagaaaaagggcgagaaagatctatt 544
 Db 496 ATGACCTTCCAGCCACAAATTTGACTTTCATTGTAAAGGAACCTGGACAGAAAGCTGACT 555
 QY 545 atgctgggtcttccaggggcacacacatgggctttattgcatatttccaccatgcccagc 604
 Db 556 ATGTTGGTCAATCCAGGGCACCACCATTTGTTTATTGCTTTTCTACAAATCCCAAGC 615
 QY 605 tggcctcagaaaaatcaaatgtatttgccttagcaccatagcactggttaagcatcaa 664
 Db 616 TGCTGAAGAATCAAAACCTTTTATGCATTAGCTCCAGTCGCCACCTGTGAAGTATACAA 675
 QY 665 aaagcccgagggaacaaattttgtgtgcagatatgatgaacaaaggattgttttggca 724
 Db 676 AAAGCTTGTAAACAAACTTAGGTTTATTCTCCAACTGTTCAAGATATATATTGGTG 735
 QY 725 aaaaagaattctgtatcagaccagatttctcagacacactgttatttacttctgtggcc 784
 Db 736 ACAAAATATTCTACCCACACAATTTCTTGTATCAATTTCTTGCCACTCAAGTGTGTCCC 795
 QY 785 aggtgattcttgatcagagattttagtaaatatcatgttacttctctgggtggagaccaca 844
 Db 796 GTGAGACACTGAATGAATTTGCGCAATGCCTTATTTATCATTTTGTGGCTTTGACAGCG 855

REFERENCE 1 (bases 1 to 1365)
 AUTHORS Bodmer, M.W., Angai, S., Varranton, G.T., Harris, T.J., Lyons, A., King, D.J., Pieroni, G., Riviere, C., Verger, R. and Lowe, P.A.
 TITLE Molecular cloning of a human gastric lipase and expression of the enzyme in yeast
 JOURNAL Biochim. Biophys. Acta 909 (3), 237-244 (1987)
 MEDLINE 87299724
 COMMENT Data kindly reviewed (10-DEC-1987) by LOWE P.A.
 FEATURES
 source Location/Qualifiers
 1. .1365
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 45. .101
 /product="pot. signal peptide (AA -19 to -1)"
 45. .1241
 /codon_start=1
 /product="gastric lipase precursor"
 /protein_id="CAA29413.1"
 /db_xref="GI:758063"
 /db_xref="SWISS-PROT:P07098"
 /translation="MWLLTMSLISVLGTHGLFGLKHPGSEVTMNISQIMTYWY PNEEYEVTDGYILEVNRIPYGRKNSGNTGQRPVFLQHGLLASATNWSLNPNSL AFILADAGYVWLGNSRNTWARRNLYSPDSVEFWAFSDMAKYDLPATIDFVKK QSKQLHYVHSGQTTIGFIAPSTNPLAKRIKTFYALAPVATVYTKSLINKLREVP QSLFKFIDGKTFYHPNFDQFLATEVCSREMLNLLCSNALFICGFSKFNPTSLD VYLSNPAGTSVQNHPTQAVKSKFQAYDWGSPVQNRHMDQSQPPYINVTANVP IAVWNGKDLADPDQVGLLLPLNLIYHKEIPYNHLDFIWMADAPQEVYNDIVSM ISEDKK"
 63. .1241
 /codon_start=1
 /product="gastric lipase precursor"
 /protein_id="CAA29414.1"
 /db_xref="GI:758064"
 /db_xref="SWISS-PROT:P07098"
 /translation="MASLISVLGTHGLFGLKHPGSEVTMNISQIMTYWYYPNEEY AYTDEGYILEVNRIPYGRKNSGNTGQRPVFLQHGLLASATNWSLNPNSLAFILAD AGYVWLGNSRNTWARRNLYSPDSVEFWAFSDMAKYDLPATIDFVKKQSKQL HYVHSGQTTIGFIAPSTNPLAKRIKTFYALAPVATVYTKSLINKLREVPQSLFK FIDGKTFYHPNFDQFLATEVCSREMLNLLCSNALFICGFSKFNPTSLD VYLS NPAGTSVQNHPTQAVKSKFQAYDWGSPVQNRHMDQSQPPYINVTANVP IAVWNG KDLADPDQVGLLLPLNLIYHKEIPYNHLDFIWMADAPQEVYNDIVSM ISEDKK"
 63. .101
 /product="pot. signal peptide (AA -13 to -1)"
 102. .1238
 /product="gastric lipase (AA 1-379)"
 BASE COUNT 396 a 284 c 268 g 417 t
 ORIGIN
 sig_peptide 63. .101
 mat_peptide 102. .1238
 Query Match 37.0%; Score 470; DB 93; Length 1365;
 Best Local Similarity 64.2%; Pred. No. 7.3e-129;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 QY 119 ctgtgaccagagcattcatgaataattagtgaaatcatcaacatcaagctatccct 178
 Db 121 CTGGAAGCCCTGAAGTACTATGACATGAGTATGATGATGATGATGATGATGAT 180
 QY 179 gtgaggaatatgaagtcgcaactgaagatgggtatatactcttctgttaacaggatccctc 238
 Db 181 ATGAAGATATGAAGTTGTGACTGAGATGGTTATATTTCTTGAGTCAATAGATTCCTT 240
 QY 239 gaggcctagtgaacctagagacaggttccaggcctgtgtgttactgcagcagctggcc 298
 Db 241 ATGGGAAGAAAAATTCAGGGAATACAGGCCAGAGACCTGTGTGTGTTTTCAGCATGGTT 300
 QY 299 tagttgaggtgtctagcaactgattcccaactgcgcacacaaatgacctggtgttctc 358
 Db 301 TGCCTGCATCAGCCCAAACTGATTTCCAACTGCCGCAACAGCAGCTTGCCTTCATTC 360
 QY 359 tggcagatgctgtgtttgtgagtggtgaggaacagcaggggaaacgcctgtgtctcgaa 418
 Db 361 TGGCAGATGCTGGTTATGATGTGTGCTGGGCAACAGCAGAGAAACCTTGGGCCAGAA 420

QY 419 aaacaagacactctccatagaccagaagatgattctgggcttccagtattgatgagatgg 478
 Db 421 GAAACTTGTACTATTACCACAGATTCAGTTGAATCTGGGCTTTTCAGCTTTTGAATGG 480
 QY 479 ctaggttgacctctctgcagtgatataaactttttttgcagaaaacggcgcaagaaaga 538
 Db 481 CTAATATGACCTTCCAGCCACAACTGACTTCATTGTTAAAGAACTGGACAGAGCAGC 540
 QY 539 tctattatgtcggctacttcacaggccaccacatgggtcttattgattttccaccatgc 598
 Db 541 TACACTATGTTGGCCATTCAGGAGCCACCACTTGGTTTATGTCCTTTTCCACCAATC 600
 QY 599 cagagctggctcagaaaatacaaatgtatttcttagcaccatagcaccactgttaagc 658
 Db 601 CCAGCCTGGCTAAAGAAATCAAAACCTTCTATGCTCTAGCTGTGTGTGTCACATGTGAAGT 660
 QY 659 atgcataaagccccggggaccataattttgtgtgcagatatgatgatacaagggatgtg 718
 Db 661 ATACAAAAAGCCCTTATAAACAACCTTAGATTCTTCTCAATCCCTCTTCAAGTTTATAT 720
 QY 719 ttggcaaaaaagaattctgtatcacagcagattctctcagacaactgttattaccttt 778
 Db 721 TTGGTCACAAATATTCACCCACACAACTTCTTGTATCAATTCCTGCTACTGAAGTGT 780
 QY 779 gtggccaggtgattcttctgtatcagattgttagtaatatcatgttactctctgggtgattca 838
 Db 781 GCTCCGTCGAGATGCTGAATCTCCTTTGCAGCAATGCCTTATTATTAATTTGTGATTTG 840
 QY 839 acacaacaatatgaacatgagccgagcagtgatatgctgccacactcttctgtggaa 898
 Db 841 ACAGTAAGAACTTAAACAGCAGTCGCTTGGATGTGTATCATCATATAATCCAGCAGAA 900
 QY 899 catctgtgcataatatctacatgagccagcagtgaaattctgtggaactctcgccgcat 958
 Db 901 CTTCTGTTTCAAAACATGTTCCATTGGACCCAGGCTGTTAAGTCTGGGAAATTCGAAGCTT 960
 QY 959 ttgactgggggagtgagacaaaatactggaaaatgcaatcagcaactcctgtaaggt 1018
 Db 961 ATGACTGGGGAAGCCAGTTCAGAAATAGGATGACATGATGATGATGATGATGATGATGAT 1020
 QY 1019 acagagtcagagatatgcaggttccctcacagcaatgtggcagaggtcagagctggttt 1078
 Db 1021 ACAATGTGACAGCCATGAATGTACCAATGTGAGTGTGGAACGGTGGCAAGGACCTGTTGG 1080
 QY 1079 caaatccagaacagtgaaaagctctctctgtgagtgaccacacatcatcaccataaga 1138
 Db 1081 CTGACCCCCAAGATGTGGCCTTTTGTCTCCAAAATCCCAATCTATTATTACCAAGG 1140
 QY 1139 ataticctgaatgggctcagtgatttctatctgggggtttggatgctcctcacgctatgt 1198
 Db 1141 AGATTCTTTTACATCATCTTGGACTTATCTGGCAATGGATGGCCCTCAGAAAGTTT 1200
 QY 1199 acaatgaaatcatcctcatctgat 1220
 Db 1201 ACAATGACATTGTTCTATGAT 1222

RESULT 8

A01046
 LOCUS A01046 1367 bp mRNA PAT 16-MAR-1993
 DEFINITION H.sapiens mRNA for human gastric lipase.
 ACCESSION A01046
 VERSION A01046.1 GI:34241
 KEYWORDS gastric lipase.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1367)
 AUTHORS POLYPEPTIDE AND POLYPEPTIDE COMPOSITION
 TITLE Patent: WO 8601532-A 5 13-MAR-1986;
 JOURNAL

Db	723	TTGGTGACAAAATATCTACCCACACAACTCTTTTGATCAATTTCTTCTACTGAAGTGT	782
Qy	779	gtgcccaggtgattcttgatcagatttgtagtaatatcatcttacttctgggtgagttca	838
Db	783	GCTCCGTCGAGATGCTGAATCTCCTTTCAGCAATGCCCTATTTTATAATTTGTGGATTG	842
Qy	839	acaccaacaatatgaatcagcagcagcaagtgtatatgtcgcacacactcttctgtaa	898
Db	843	ACAGTAAGAACCTTAACACGAGTCGCTTGGATGTGTATCTATCATATATCCAGCAGAA	902
Qy	899	catctgtcaaaaattcttcaactcagcagcagcagtaaatcttgtaactcccgccat	958
Db	903	CTTCTGTTCAAAACATGTTCCATTGGACCCAGCTGTTAAGTCTGGGAAATTTCCAAGCTT	962
Qy	959	ttagctgggggagtgagacccaaaatctggaaaatgcaatcagccaaactcctgaegt	1018
Db	963	ATGACTGGGGAAGCCAGTTCAGATAGGATGACATCATCATCAGTCCCAACCTCCCTACT	1022
Qy	1019	acagagtcagagatagcaggtccctacagcaatgtgacacagaggtcagagctgcttt	1078
Db	1023	ACAATGTGACAGCCATGAATGTACCAATTGCAATGTGGAACGGTGGCAAGACCTGTTGG	1082
Qy	1079	caaatccagaagacgtgaaatgtcctctcaggtgaccaaactcattaccataaga	1138
Db	1083	CTGACCCCCAAGATGTTGGCCTTTTGCCTTTCCAAAACCTCCCCAATCTTATTACCAAGG	1142
Qy	1139	atattctgaatgggctcacgtgatttcatctggttggattgctcctcaccgtatgt	1198
Db	1143	AGATTCCTTTTACATCATCTTGGACTTTATCTGGCAATGATGCCCTCAAGAAGTTT	1202
Qy	1199	acaatgaaatcatccatctgat	1220
Db	1203	ACAATGACATTGTTCTATGAT	1224
RESULT	9		
AL2714	LOCUS	1367 bp	DNA
AL2714	DEFINITION	Pregastric lipase.	
AL2714	ACCESSION	AL2714	
AL2714.1	VERSION	GI:579560	
human.	KEYWORDS		
Homo sapiens	ORGANISM		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
1 (bases 1 to 1367)	REFERENCE		
PRECURSOR POLYPEPTIDE, DNA SEQUENCE CODING THEREFOR, VECTORS, HOST	AUTHORS		
ORGANISMS, AND PROCESSES INVOLVING SAME	TITLE		
Patent: WO 8603778-A 8 03-JUL-1986;	JOURNAL		
Location/Qualifiers	FEATURES		
1. .1367	source		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/codon_start=1			
/product="pregastric lipase"			
/protein_id="CAA01053.1"			
/db_xref="GI:579561"			
/db_xref="SWISS-PROT:P07098"			
/translation="MWLLTMSLISVLGTHGLFGLKLPSPVETMNIQSOMITWY			
PNEYEVVTEDEGYLEVNRIPYKKNNGTQRPVVFLOHGLLASATNWSINLPNLS			
AFILADAGYVHSGNNTWARRNLYSPDSVEFWAFSDMAKYDLPATDFIVKK			
TQKOLHYVGHSGGTTIGTAFSTNPISLAKRIKTFYALAPVATVYKTSILNKLREVP			
QSLFKFIFGDKIPYHNFPDQFLATEVCSEMLNLLCSNALFICGFDKSNFTSRLD			
VYLSHPACTSVQNMHFHTQAVKSGKFOAYDNGSPVQNMHYDQSPPYNYVTAMNVP			
IYVNGGKDLADPDQVGLLLPKLPNIYHKEIPYHNHLDFIWAMDAPQEVYNDIVSM			
ISDDK"			
397 a	284 c	269 g	417 t
BASE COUNT			
ORIGIN			
Query Match 37.0%; Score 470; DB 9; Length 1367;			
Best Local Similarity 64.2%; Pred. No. 7.3e-129;			
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;			
Qy	119	ctgtggaccagagcattcatgaattattagtgaaatcatcccaacatcaagggtatccct	178
Db	123	CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCAA	182
Qy	179	gtagggaatgaagtcgaactgaagatgggtatctcttctgttaacagatccctc	238
Db	183	ATGAAGATATGAAGTTGTGACTGAAGATGTTATATCTTGAAGTCAATAGAATCCCT	242
Qy	239	gaggcctagtgcacctaagaagacaggttccagggcctgtggtgttactcagcagtgcc	298
Db	243	ATGGGAAGAAATTCAGGGAATACAGCCGACAGACCTGTGTGTTTTCACAGCTGTT	302
Qy	299	tagttgaggtgctagcaactgatttcccaactgcgccaacatagcctggggttcattc	358
Db	303	TGCTTGATCAGGCACCAAACTGGATTTCCAACTGCCGGAACACAGCAGCTTGCCTTATC	362
Qy	359	tggcagatgctgttttgacgtgtgatgggaacagcaggggaacacgctgtctcgaa	418
Db	363	TGGCAGATGCTGGTTATGATGTGTGGCTGGGCAACACAGAGGAAACACCTGGGCCAGAA	422
Qy	419	aacacaagacactctccatagaccagaatgagttctgtgggtttcagttatgatgagatg	478
Db	423	GAACTTGACTATTACAGATTCAGTTCAATTCGGGCTTTCAGCTTTGATGAATGG	482
Qy	479	ctaggtttgaccttcctcgcagtgataaactttattttgcagaaaaagcggccagaaaga	538
Db	483	CTAAATATGACCTTCCAGCCACCAATCGACTTCATTTGTAAGAAAACCTGGACAGAGCAGC	542
Qy	539	tctattatgctgctattcaacagggaacacattgggtcttattgattttccaccatgc	598
Db	543	TACACTATGTTGGCCATTCACGGGACCACTTGTGTTTATGGCTTTTCCACCAATC	602
Qy	599	cagagctggctcagaaaaatcaaatgtattttgcttttagcaccatagccactgttaagc	658
Db	603	CCAGCTTGGCTTAAAGAAATCAAACTCTATGCTCTAGCTCTGTTGGCAGCTGGAAGT	662
Qy	659	atgcaaaaagcccgaggaccataattttgtgtgctgcagatgatgatcaagggttgt	718
Db	663	ATACAAAAGCCTTATAAACAAACTTAGATTGTTGCTCAATCCCTCTCAAGTTTATAT	722
Qy	719	ttggcaaaaagaattctctgatcagacacagatttctcagacaaactgttattacccctt	778

Db 693 GTTTGGACAGCAATTTCTCCCTCAGAGTGCATGCTGAAGTGGCTGACATTCACGT 752
 Qy 777 ttgtgcccagcgtgattcttgatcagattgttagtaataatctgttactctctggtgatt 836
 Db 753 TTGACCCACGCTCATCATGAAGAGCTTTGGCAACGCTCTCTCTGCTATGCTGCTT 812
 Qy 837 caacacacacatgaacatgagccagcagcaagtgatgatatctgctccacactctgctg 896
 Db 813 CAATGAGAGAGATTTAAATGCTAGAGTGGATGTGTATACACACACTGCCCTGGCGA 872
 Qy 897 aacatctgtgcacaaatattctacactgagccagcagcagtggaattctgtgaactccggc 956
 Db 873 ACTTCTGTCGCAAAACATGTTACACTGGGTCAGGTTTAAATACCGTAGCTTCAAGC 932
 Qy 957 atttgactgggggagtgagacacaaatctggaaaaatgcaatcagcagcaactcctgtaag 1016
 Db 933 TTTTACTGGGGAAGCAGTGAAGAGAAATTTATTCATTACACACAGAGTTTTCCTCCTC 992
 Qy 1017 gtacagagtcagagatgacggtccctacagcaaatgtgacagagagtcaggactggct 1076
 Db 993 GTACAAACATAAAACATGCTGTACCCACTGCTTTGGAGCGGGCGGGGACTGGCT 1052
 Qy 1077 ttcaaatccagagacgctgaaatgctgctctgaggtgacacacactcatctaccataa 1136
 Db 1053 GGCAGACATCAATGACATCACTATCTTACTGACATCAATCCCAAGTTGGTGTATCAAA 1112
 Qy 1137 gaattactctgaatgggctcagctggtatctctggtggttggatgctctcaccgtat 1196
 Db 1113 GAACATTCCTGAGTGGGACCATCTGGATTTTATCTGGGTTTGGATGCCCTCGGAACT 1172
 Qy 1197 gtacaatgaatatcatctcatctgagcagcagggag 1233
 Db 1173 GTATGATGAATAATAGTCTAATGAAGAAATACCAG 1209
 RESULT 13
 S81497
 LOCUS 3144 bp mRNA ROD 02-AUG-2000
 DEFINITION lysosomal acid lipase-intracellular hydrolase [rats, Wolman, liver,
 mRNA, 3144 nt].
 ACCESSION S81497
 VERSION S81497.1 GI:1336725
 KEYWORDS Rattus sp. liver Wolman.
 SOURCE Rattus sp.
 ORGANISM Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 3144)
 AUTHORS Nakagawa,H., Matsubara,S., Kuriyama,M., Yoshidome,H., Fujiyama,J.,
 Yoshida,H. and Osame,M.
 TITLE Cloning of rat lysosomal acid lipase cDNA and identification of the
 mutation in the rat model of Wolman's disease
 J. Lipid Res. 36 (10), 2212-2218 (1995)
 JOURNAL 96129534
 MEDLINE GenBank staff at the National Library of Medicine created this
 entry [NCBI gibbs 176569] from the original journal article.
 REMARK This sequence comes from Fig. 2.
 FEATURES
 source Location/Qualifiers
 1..3144
 /organism="Rattus sp."
 /db_xref="taxon:10118"
 1..3144
 /gene="lysosomal acid lipase, LAL"
 233..1426
 /gene="lysosomal acid lipase, LAL"
 /note="intracellular hydrolase; This sequence comes from
 Fig. 2; conceptual translation presented here differs from
 translation in publication; LAL"
 /codon_start=1
 /product="lysosomal acid lipase"
 /protein_id="A836043.2"
 /db_xref="GI:9653291".

/translation="MQLLGRVICFVVGILLSSGGTGTISAVDPNANNVTIIMHWGY
 PHSVDTGDIILGVHPRKQNFQDKPKPVYVILQMRHGFADSNWYTNIDNNSL
 GFTLADAGFDVWNGSRNKTLSVQSDYWFSDYMAKVPKGLRLP
 TGOEQLNVHSGOCTIGFTAFSOMPELAKKVMFFALAPVLSLNFASGPMVKGLRLP
 DLLEDLFGOKQFLOPSAMVKKWLSHICTHVINKELCANIFFLLICGFENKLNLSRVD
 VYTHCPAGTSVONVHWTOVKYKQLQAFDWSSDKNYFHYNOSYPLYSIKDKQLP
 TALWSGKDWLADTSDINILTEPLTVLHKHIFEWHDHLDIFMGDAPWRDLNEYVSL
 MKYIQ"
 BASE COUNT 852 a 690 c 754 g 848 t
 ORIGIN
 Query Match 36.2%; Score 459.8; DB 95; Length 3144;
 Best Local Similarity 64.5%; Pred. No. 9e-126;
 Matches 724; Conservative 0; Mismatches 387; Indels 12; Gaps 2;
 Qy 117 agctgtgaccagaagcattcatgaattattagtgaaatcatccaacatcaaggtatccc 176
 Db 307 AGCTGTGACCTGAAGCAACATGAATGTGACTGAGATAATCATGCTGGGGATATCC 366
 Qy 177 ctgtgaggaatatgaagtcgcaactgaagatgggtatatactcttctgttaacagagattcc 236
 Db 367 A-----GAGCACTCAGTACAGACAGGAGATGGCTACATTCTGGGTGTTACCGAATCCC 420
 Qy 237 tcgagggcctagtcaacctgaagaagacaggttccaggcctgtgtgttactgcag----- 291
 Db 421 GCATGGCGGAAGAACAGATTTGATAAGGTCCCAACACAGTTGTGTATCTTCAGTGGCG 480
 Qy 292 -catggcctagtggaggtgctagcaactgagatttccaaactgcccacacatagcctggg 350
 Db 481 ACATGGCTTCTGCGCAGATCTTAGTAACCTGGGTGTTGATGGGAAACACAGACACACCTGG 540
 Qy 351 cttcattctgagcagatgctgttgcagtgatgggggaacagcagcaggggaacacgctg 410
 Db 541 CTTTCATCTGGCAGATGCTGGCTTTGATGTGTGATGGGAAACACAGACAGAAACACCTG 600
 Qy 411 gtctgaaaaacacagacactctccatgacacaaagatgagttcttggggttccagtattga 470
 Db 601 GTCTCGGAAGCACAAAGACTCTGTCAAGTTCCTCAGATGAATACACAGGCTTCAGTTTGA 660
 Qy 471 tgagatggctaggtttgacctctcagtgatgataaaacttttttgcagaaaaacgggcca 530
 Db 661 TGAGATGGCAAAATATGACCTCCCTGCTCAATTAACATACATACATTTTGAATAAACTGGCA 720
 Qy 531 ggaagaatctattatgctcgctattccagcagggcaccacacagccttattgacttttc 590
 Db 721 AGAACACCTGTATTAATGTGGCCCACTCTCAAGGCTGCACCATAGTGTTTTATAGCATTTTC 780
 Qy 591 caccatgccagagctggtctcagaaaaatcaaaatgtatttttcttagcaccatagccac 650
 Db 781 ACAGATGCTGAACCTGGCAAGAGAGTTAAATGTTTTTCCCTGGCTCCTGTGTTGTC 840
 Qy 651 tgttaagcatgcaaaaagccccgggacaaaatattttgtgtgcagataatgatgataa 710
 Db 841 TCTCAATTTTGTCTCGGCGCGATGGTCAAAATGGGAGGTTGCGGATCTGCTTCTTGA 900
 Qy 711 gggatttttggcaaaaagaattctgtatcagacagatcttctcagaaactgttat 770
 Db 901 GGACTTGTGTGGGCGAAGCAATTTCTTCTCAGATGCAATGGTGAATGGCTAAGCAC 960
 Qy 771 ttacctttgtggccaggtgattcttgcagattgttagtaataatcattgttactctggg 830
 Db 961 TCACATTTGCGACGACGCTCATCATCAGGAGCTTTGCCCAACATCTTCTTCTGATATG 1020
 Qy 831 tggattcaacacacacaaatatagaacatgagccgagcagtgatgatactgccccactct 890
 Db 1021 TGGTTCATAGAGAAGAAATTTAAACATGCTCTAGAGTGGATGTGTATACAAACACACTGTC 1080
 Qy 891 tgcgtggaacatctgtcaaaatattctacactggagccagcagtggaattctggtgaact 950
 Db 1081 TGGCGGAACCTCTGTGTCAAAACATGGTACACTGACCCAGGTTGTTAAATACCATAAGCT 1140
 Qy 951 ccgggcatctgactgggggagtgagacacaaaatctggaaaaatcgaatcagcagcactcc 1010

Db 1141 TCAAGCTTTTACAGTGGGGAACAGTACAGAAATATTTTTCATTACAAATCAGAGTTATCC 1200
 Qy 1011 tgytaagggtacagagtcagagatgacggtccctacacagcaatgtggacagagagtcagga 1070
 Db 1201 ACCCTTGTACAGCATAAAGACATGCAACTACCCACCTGCTTTATGGAGCGGGGCAAGGA 1260
 Qy 1071 ctggcttcaaatccagaagacgctgaaatgctgctctctctgaggtgacacacccatccta 1130
 Db 1261 CTGGCTGGCAGACACACAGTGCATCAATATCTTACTGACTGAATCCCCAGTTAGTGA 1320
 Qy 1131 ccataagaataattcctgaatgggtcagctgagattcctcctgaggtgaggtgaggtcctccta 1190
 Db 1321 CCACAAGAACATCCCGAGTGGGACCATCTGGATTTATCTGGGGTTTGGATGCCCGCTG 1380
 Qy 1191 ccgtatgtacatgaatccatccatctgagtcagcagcagagagag 1233
 Db 1381 GAGGCTGTATTAAGTAGTACGTCTTANTGAAGAAATACCAG 1423

RESULT 14

BOVME BOVME 1417 bp mRNA MAM 10-DEC-1994
 LOCUS Bos taurus pregastric esterase mRNA, complete cds.
 DEFINITION L26319
 ACCESSION L26319
 VERSION L26319.1 GI:600756
 KEYWORDS esterase; pregastric esterase.
 SOURCE Bos taurus 25 weeks Tongue serous gland cDNA to mRNA.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 1417)
 TImmermans,M.Y.J., Kuipers,L.P. and Teuchy,H.
 The cDNA sequence encoding bovine pregastric esterase
 JOURNAL Gene 147, 259-262 (1994)
 MEDLINE 95011625

FEATURES

source
 1. .1417
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /cell_type="Secretory cell"
 /dev_stage="25 weeks"
 /tissue_type="Tongue serous gland"
 73. .129
 /EC_number="3.1.1.3"
 /product="pregastric esterase"
 /protein_id="AA57037.1"
 /db_xref="GI:600757"
 /translation="MWWLLVTVCFIHMSGNFCLFKIAKNPEASMNVSOMISYWGYP
 SEMHKVITADGVILOVYRPHGKNANHLGQPVVFLQHLGSGATNWSLNPKNSLG
 FLILDAGYDVWLGNSRGNTWAOEHLVYSPDSEFWAFSEDAEYDLPTIDFILRLP
 GOKLHYVCHSOGTTIGFIATFSTPLAEKIKVYALAPVATVYKTKSLFNKLALIPH
 FLEKIFGDKMTPIHFLQFLGVEMCSRETLVLCNKALFAITGVDNKFNMSRLDY
 YIAHNPAGTSVONTLHWRAVSKGFQAFDQAPYQNLNMYHQPTPIYNLFAMNVPY
 AYSADNDLLADPDQVDFLLSKLSNLNLYHKEIPNYNHLDFIWMADAPQBYNIVSLM
 AEDKK"
 130. .1263
 /EC_number="3.1.1.3"
 /function="triacylglycerol acylhydrolase"
 /evidence="experimental"
 /product="pregastric esterase"
 1378. .1385
 polyA_signal
 1396
 polyA_site
 408 a 306 c 276 g 427 t
 BASE COUNT
 ORIGIN

Query Match 36.1%; Score 458; DB 7; Length 1417;
 Best Local Similarity 62.1%; Pred. No. 2.8e-125;
 Matches 741; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

Qy 44 tggaaatgtggcttctgattctggtggttatattgttccagagaaatgtgaattcagtaac 103
 Db 68 TGAATAATGTGGTGGCTACTTGTAAACAGTGTGTTTCATCCACATGCTCTGGAATATGCATTTT 127
 Qy 104 atatgc---caactaaagctgtgcccagagcattcatcatgaatattagtgaaacatccc 160
 Db 128 GTTTCTTGGAAATAATGCTAAGAACCTGAAGCCAGTATGAATGTTAGTCAGATGATTT 187
 Qy 161 aacataagcctatccctgtgaggaatgatgaagtgcgaactggaagatgggttatcccttt 220
 Db 188 CTTACTGGGGCTACCCAAAGTGAGATGCATAAAGTTAATCACTGCGATGCTTATATCCCTTC 247
 Qy 221 ctgttaacagagattccctcagagccctagtgaacctaagaagacaggttccagggcctgtgg 280
 Db 248 AGGCTATCGGATTCCTCATGGAAGAATAATGCTAATCATTTAGGTCAGAGACCTGTGTG 307
 Qy 281 tgttactgcagcatggcctagtgtgaggtgctagcaactggatttccaaacctgcccaaca 340
 Db 308 TGTTCCTGCAGCATGGTCTTCTTGGATCAGCCACAACTGGATTTCCAACCTGCCCAAGA 367
 Qy 341 atagcctgggcttccattctctggcagatgctggttttgacgtgtgagtgaggggaacagcggg 400
 Db 368 ACAGCTGGGCTTCTCTCTGGCAGATGCTGTTATGACGTGTGGTGGGGAACAGCAGAG 427
 Qy 401 gaaacgctggtctcggaaacacacacacactctccatagacacagagatgagttctgggctt 460
 Db 428 GAAACACCTGGGCCAGGAACATTTATCTATTTACCAGACTCCCGGAATCTTGGGGCTT 487
 Qy 461 tcsagtattgatgatgagtggttaggtttgacctctctgctgagtaaaactttattttgcaga 520
 Db 488 TCAGCTTTGTATGAATGGCGGAATATGACCTTCCATCTACAAATGATTTTCATCTAAGGA 547
 Qy 521 aaacgggccaaggaagaatctattatgctggtattcacaggggcacacacatgggcttta 580
 Db 548 GAACAGGACAGAAGAGCTACATATGTGGCCATTGCCAAGGCACCACTTTGGTTTGA 607
 Qy 581 ttgcattttccaccatgccagagcgtggtctcagaaaaatcctgataaaactttatttgcctacac 640
 Db 608 TCGCTTTTCTACAGTCCCATGCTGGCTGAAAAAATCAAAAGTCTTCTATGCATTAGCCC 667
 Qy 641 ccatagccactgttaagcatgcaaaaagccccgggagcccaaatattttgttgcgtccagata 700
 Db 668 CAGTTGCCACAGTGAAGTACACCAAGAGCCTGTTTAAACAAACTTGCATTTATTCCTCACT 727
 Qy 701 tgaatgataagagatgtgttggcaaaaaaagaattctctgatacagacagattctccagac 760
 Db 728 TCCTCTTCAAGATTATATTGGTGACAAAATGTTCTTACCACACACTTTTGTGGAACAAT 787
 Qy 761 aactgttattacccttggccaggtgattcttgcagatttgcagtaataatcatgt 820
 Db 788 TTCTTGGTGTGAATGTGCTCCGTGAGACACTGGATGCTCTTGTGAAGATGCCCTGT 847
 Qy 821 tacttctgggtgattcaac 880
 Db 848 TTGCCATTACTGGAGTGAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 907
 Qy 881 cccacactctgtggaacatctgtgcaaaaatattctacactggagcagcagcagcagcagcagc 940
 Db 908 CACATAATCCAGCAGGAACCTTCTGTTCAAAACACCCCTCCACTGGAGACAGGCTCTTAAGT 967
 Qy 941 ctggtgaaactccgggcatcttgactgggggagtgagacacacacacacacacacacacacacac 1000
 Db 968 CTGGGAAATTCCAAGCTTTTGACTGGGAGCCCCCATATCAGAACCTTAATGCATTATCATC 1027
 Qy 1001 agccaactcctgtaaggtacagagtcagagatatgacggtccctcacagcaatgtggacag 1060
 Db 1028 AGCCACACCTCCCATCTACATAATTTAACAGCCATGAATGTCCTCAATTCGATATGGAGTG 1087
 Qy 1061 gaggtcagagactggcttccaaatccagagaagacgtgaaaaatgctgctctcagagtgacca 1120
 Db 1088 CTGAAATACCTGTGGGTGACCCCTCAGGATGTTGACTTTTCTGCTTTTCAAAACTCTCTA 1147

GenCore version 4.5
Copyright (C) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:41:42 ; Search time 211.8 Seconds
(Without alignments)
3762.073 Million cell updates/sec

Title: US-09-333-159-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtcaag.....gacggtgtgagccggtattg 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_0601:*
- 1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
 - 2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
 - 3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
 - 4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
 - 5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
 - 6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
 - 7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
 - 8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
 - 9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
 - 10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
 - 11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
 - 12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
 - 13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
 - 14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
 - 15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
 - 16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
 - 17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
 - 18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
 - 19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
 - 20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
 - 21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
 - 22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1269	22	AAF45132 Human TANGO 294 OR
2	1269	100.0	2044	22	AAF45131 Human TANGO 294 CD
3	489.2	38.6	1378	14	AAQ42310 RGL precursor. Or
4	470	37.0	1367	7	AAQ60685 Sequence encoding
5	470	37.0	1367	7	AAQ60566 Sequence encoding
6	470	37.0	1367	17	AAQ58916 Human gastric lipa
7	464.4	36.6	1336	6	AAQ50385 Rat prelingual lip
8	444.2	35.0	1528	15	AAQ68388 Canine gastric lip
9	444.2	35.0	1528	17	AAQ58915 Dog gastric lipase
10	444.2	35.0	1531	15	AAQ68389 Canine gastric lip
11	286	22.5	699	22	AAF28679 Human protein HP03

12	285	22.5	1308	22	AAF28689 Human protein HP03
13	257.3	20.3	617	21	AAA44349 Human secreted exp
14	193.2	15.2	572	21	AAA43744 Mouse secreted exp
15	100	7.9	1718	21	AAZ34958 Soybean acid triac
16	89.4	7.0	1483	21	AAZ34956 Rice acid triacylg
17	84.8	6.7	859	21	AAZ34953 Corn acid triacylg
18	72	5.7	1438	21	AAZ34959 Soybean acid triac
19	68	5.4	936	22	AAF58252 Oligonucleotide D1
20	68	5.4	936	22	AAF58254 Oligonucleotide D1
21	68	5.4	936	22	AAF58257 Oligonucleotide D1
22	68	5.4	936	22	AAF58259 Oligonucleotide D2
23	68	5.4	936	22	AAF58262 Oligonucleotide D2
24	68	5.4	938	22	AAF58255 Oligonucleotide D1
25	66.2	5.2	936	22	AAF58252 Oligonucleotide D1
26	66.2	5.2	936	22	AAF58254 Oligonucleotide D1
27	66.2	5.2	936	22	AAF58257 Oligonucleotide D1
28	66.2	5.2	936	22	AAF58259 Oligonucleotide D2
29	66.2	5.2	936	22	AAF58262 Oligonucleotide D2
30	66.2	5.2	938	22	AAF58255 Oligonucleotide D1
31	55.8	4.4	705	21	AAZ34952 Catalpa acid triac
32	54	4.3	549	21	AAC93827 Cat flea hindgut a
33	51	4.0	509	21	AAZ34954 Cell death regulat
34	48.2	3.8	178	19	AAV35431 Nucleotide sequenc
35	46.4	3.7	43676	21	AAZ60606 Methanococcus jann
36	38.6	3.0	1664976	19	AAV21209 Complete nucleotid
37	36.8	2.9	13254	21	AAA32165 Oligonucleotide D1
38	36.6	2.9	244	22	AAF58238 Human G-protein co
39	36.6	2.9	3271	19	AAV00060 DNA encoding VH re
40	36.6	2.9	3282	17	AAT40914 Modified heavy cha
41	36.6	2.9	3282	21	AAA31024 BL2 IgG1 preferred
42	36.6	2.9	3282	21	AAA31038 BL2 IgG1 preferred
43	36.6	2.9	3282	21	AAA32150 Modified heavy cha
44	36.6	2.9	3282	21	AAA32164 BL2 IgG1 preferred
45	36.6	2.9	5412	20	AAZ21356 Human BAI2 gene.

ALIGNMENTS

RESULT 1

AAF45132
ID AAF45132 standard; cDNA; 1269 BP.
XX
AC AAF45132;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 ORF.
XX

KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.

XX Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX P-PSDB; AAB66065.

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease

Claim 1; Fig 6; 359pp; English.

The present invention relates to RANGO or INTERCEPT sequences (see AAF45121-F45136 and AAF45137-F45150, AAB66031-B66057, AAB66064-B66083 and AAB66085). The RANGO/INTERCEPT proteins and coding sequences are useful for the treatment of neurological disorders such as central nervous system (CNS) disorders, CNS-related disorders, focal brain disorders, global-diffuse cerebral disorders and other neurological and cerebrovascular disorders. The CNS disorders include Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, autonomic function disorders such as hypertension and sleep disorders, neuropsychiatric disorders, psychoactive substance use disorders, anxiety, and bipolar affective disorder.

Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

```
Query Match      100.0%; Score 1269; DB 22; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	atgttggaacacctgttcaagacacagtggtgattgtctcacacagaatggaaatgtggtctctg	60
Db	1	atgttggaacacctgttcaagacagtggtattgtctcacacagaatggaaatgtggtctctg	60
QY	61	attctggtggcgtatatgtttccagagaaatgtgaaattcagtgacatatgccaactaaagct	120
Db	61	attctggtggcgtatatgtttccagagaaatgtgaaattcagtgacatatgccaactaaagct	120
QY	121	gtggaccagaagcatttcagataattagtgaataatcatccaacatcaaggctatccctgt	180
Db	121	gtggaccagaagcatttcagataattagtgaataatcatccaacatcaaggctatccctgt	180
QY	181	gaggaataagaatgcgaactgaagatgggtatatatcttctgttaacaggattccctga	240
Db	181	gaggaataagaatgcgaactgaagatgggtatatatcttctgttaacaggattccctga	240
QY	241	ggcctagtgcacactaaagaagacaggttccaggcctgtgtgttactgcagatggccta	300
Db	241	ggcctagtgcacactaaagaagacaggttccaggcctgtgtgttactgcagatggccta	300
QY	301	gttggagggtgtgtagcaactggatttccaacctgtcccaacaatagcctgggttcattctg	360
Db	301	gttggagggtgtgtagcaactggatttccaacctgtcccaacaatagcctgggttcattctg	360
QY	361	gcagatgcctgttttccactgtgtgagtggggaacacagcaggggaaacccctgtctcgaaaa	420
Db	361	gcagatgcctgttttccactgtgtgagtggggaacacagcaggggaaacccctgtctcgaaaa	420
QY	421	cacaagacactctccatagaccagaatgagttctggccttccagttatgatagagatggct	480
Db	421	cacaagacactctccatagaccagaatgagttctggccttccagttatgatagagatggct	480
QY	481	aggtttgacctctcctgcagtgtataaactttatttgcagaaaaacggccaggaaaaagatc	540
Db	481	aggtttgacctctcctgcagtgtataaactttatttgcagaaaaacggccaggaaaaagatc	540
QY	541	tattatgtggcgtattcacagggccacacacattggtttattgcattttccaccatgccca	600
Db	541	tattatgtggcgtattcacagggccacacacattggtttattgcattttccaccatgccca	600
QY	601	gagctggctcagaaaaatcaaaatgtatttgccttagcaccataagccactgttaagcat	660
Db	601	gagctggctcagaaaaatcaaaatgtatttgccttagcaccataagccactgttaagcat	660
QY	661	gcaaaagcccccggaccaaaattttgtgtgtgccagatatgatgatcaaggagattgttt	720

Db	661	gcataaagcccccgggaccataatttttctgtcgcagatgatgatcaaggattgttt	720
Qy	721	ggcaaaaaagaattctgtatcagaccagattctcagacaactgttatttaactttgt	780
Db	721	ggcaaaaaagaattctgtatcagaccagattctcagacaacttgttatttaactttgt	780
Qy	781	ggcagggtgattcttgatcagatttgataataatcatcttactctgggtggattcaac	840
Db	781	ggcagggtgattcttgatcagatttgataataatcatcttactctgggtggattcaac	840
Qy	841	accnacaatatgaacatgagccgagcaagtgtatatgtgccacactcttgcctggaaca	900
Db	841	accnacaatatgaacatgagccgagcaagtgtatatgtgccacactcttgcctggaaca	900
Qy	901	tctgtcaaaaattttctacactggagccagcagtgaaattctctggtaactcgcgggcattt	960
Db	901	tctgtcaaaaattttctacactggagccagcagtgaaattctctggtaactcgcgggcattt	960
Qy	961	gactgggggagtgagaccataaattcggaaaaatgcaatcagccaaactcctgtaaggctac	1020
Db	961	gactgggggagtgagaccataaattcggaaaaatgcaatcagccaaactcctgtaaggctac	1020
Qy	1021	agactcagagatagcaggtccctacagcaatgtggacagaggtcagagactggtttcca	1080
Db	1021	agactcagagatagcaggtccctacagcaatgtggacagaggtcagagactggtttcca	1080
Qy	1081	aatccagaagacgtgaaaatgtctctctgagtgacaaacctcatctacataaagaat	1140
Db	1081	aatccagaagacgtgaaaatgtctctctgagtgacaaacctcatctacataaagaat	1140
Qy	1141	attcttcgaatgggtcacgtggaatttcattctgggtttggatgctccctcacgtagtgcac	1200
Db	1141	attcttcgaatgggtcacgtggaatttcattctgggtttggatgctccctcacgtagtgcac	1200
Qy	1201	aatgaataatccatctgtatgcagcaggaggagaccaacctttcccaggagcgtgtgag	1260
Db	1201	aatgaataatccatctgtatgcagcaggaggagaccaacctttcccaggagcgtgtgag	1260
Qy	1261	gccgatttg 1269	
Db	1261	gccgatttg 1269	
RESULT 2			
AAFA5131			
ID	AAFA5131 standard; cDNA; 2044 BP.		
XX	AAFA5131:		
XX			
XX			
DT	30-MAR-2001 (first entry)		
XX			
DE	Human TANGO 294 cDNA.		
XX			
Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder; central nervous system; focal brain disorder; bipolar affective disorder; global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease; senile dementia; Huntington's disease; amyotrophic lateral sclerosis; Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep; neuropsychiatric; psychoactive substance use; anxiety; ss.			

RESULT 2

AAAF45131
ID AAAF45131 standard; cDNA; 2044 BP

AA
AC
AAE45131:

DT 30-MAR-2001 (first entry)

XX DE Human TANGO 294 cDNA.

AA Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neurophysiologic; psychoactive substance use; anxiety; ss.

xx
os Homo sapiens.

AA
PN
WO200077239-A2.

21-DEC-2000

24-MAY-2030: 2000WO-US14858.

AA
PR 14-JUN-1939: 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM:

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM:

PR 13-NOV-1991; 91FR-0013948.
XX (LJOU) INST RECH JOUVEINAL.
XX Benicourt C, Blanchard C, Junien J;
XX WPI; 1993-161080/20.
DR P-PSDB; AAR37302.
XX Rabbit gastric lipase, its precursor and their DNA - useful for
PT treating conditions linked to gastric lipase deficiency, such as
PT mucoviscidiosis and pancreatic exocrine insufficiency
XX Claim 1; Fig 6; 31pp; French.
XX mRNA was isolated from rabbit gastric mucosal cells, converted to
CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
CC to transform E.coli MW294 and cells screened for reactivity with
CC probes corresp. to known parts of the RGL sequence. One positive
CC clone contd. plasmid p70101 which contd. a 1.35 kb PstI sequence.
CC This fragment was isolated, subjected to PCR and the amplification
CC prod. ligated with the 2.2 kb BglII-SacI fragment of pRU276 (contg.
CC a synthetic Tac promoter and transcription stop signal). The
CC recombinant plasmid was designated pRGLN2.1.
XX SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;

Query Match 38.6%; Score 489.2; DB 14; Length 1378;
Best Local Similarity 65.0%; Pred. No. 2.7e-145;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 125 acccagaagcattcatgaattagtagaatacctcaacacacagcgtatccctgtgagg 184
DB 136 accctgaagtgaattgaattagtcagatgatttctactctggtggaacccaagtga 195

QY 185 atatgaagtcgcgaactgaagatgggtatatactcttctgttaacaggattctctgagcc 244
DB 196 atatgaagttgtgactgaagatggtatatactctggttaacaggattctctgagcc 255

QY 245 tagtgcacctaagaagacaggttcacgcctggtgtgttactcagcagcgtcctagt 304
DB 256 agaaaaattcagggaacagagccagagacccgtcgtatcttctgagcagatggtgct 315

QY 305 gagggtgtagcaactggtattcccaactgcgaacacatagcctgggtcttctctggcag 364
DB 316 catcagctcaaacctggtatctcccaactgcgaacacacagcctgcttcttctggcag 375

QY 365 atgctggttttgacgtgtggtggaacagcagcaggggaacgcctggtctcgaaaaaca 424
DB 376 atgctggtttatggtgtgtggtggaacagcagcaggggaacacacctggtccagaaga 435

QY 425 agacactctccatagacccaagatgattctgggtctcagttatgattgagatgctaggt 484
DB 436 tatactattcaccagactccgtggaattctgggtctttagctttagtgaatggtcgaat 495

QY 485 ttgacctctctgcagtgataaacttttttgcagaaacggcggaagaaagatctatt 544
DB 496 atgacctctccagcccaattgactcattgttaaaggaaactggacaggaagcgtgcact 555

QY 545 atgctggtatttccagggcaccacacatggcctttattgatttttccaccatgccagagc 604
DB 556 atgctggtatttccagggcaccacacatggcctttattgatttttccaccatgccagagc 615

QY 605 tggctcagaataaataatgatttttcttagcctccatagccactgttaagatgca 664
DB 616 tggctcagaataaataatgatttttcttagcctccatagccactgttaagatgca 675

QY 665 aaagcccccgggaccac 724
DB 676 aaagcctgttaac 735

QY 725 aaaaagaattctgtatcagaccagatttctcagacacacacttgttattacaccttggcc 784

DB 736 acaaatattctaccac 795
QY 785 agggattcttgatcagatgttagtaataatcattcttctggttgattcaacaca 844
DB 796 gtgagacatgaatgtaatttgacgaatgcttatttattcttctggttgacagc 855
QY 845 acaatgaacatgagcgcgagcaagtgtatatgtctgcccacacacttctctggaacatctg 904
DB 856 caaactgaacatgagcgcgtggtgagtgtagtgatcacaataatccgcgaggaacttcag 915
QY 905 tgcacaaatattctacactggagccagcagtgaaattctctggaactccggcagcttgact 964
DB 916 ttcaaacatgctgcaactggacccagcgtgtaaatctggaacttttcaagctttaatt 975
QY 965 ggggagtgagac 1024
DB 976 ggggagtgccagcgtcagaatgtagtgcacttcaatcagccacacactccctactaca 1035
QY 1025 tcagagatgacgcgtccctacagcaatgtggacaggggtcagactggttcaaatc 1084
DB 1036 tgacgcacatgaatgccaattgcagtggtgagtggtggaactgactggttgctgacc 1095
QY 1085 cagaacagcgtgaaatgctgtctctggtggtgacacacactcattcaccataagaatctc 1144
DB 1096 cccaagatgttgacctttgtctccaaactttctaatctcatttaccacaaggagattc 1155
QY 1145 ctgaatgggtcagctgagattctctctggttgggttgggtcctcctcactgattacaatg 1204
DB 1156 ttcatataacacactggtatttctctggtggaatggtctcctcagaagattcaatg 1215
QY 1205 aaatcattcattctgagcagcagggagaga 1234
DB 1216 aaattattctctgagcagcagggagagataaaa 1245

RESULT 4
AAN60685
ID AAN60685 standard; DNA; 1367 BP.
XX AC AAN60685;
XX DT 08-JUN-1991 (first entry)
XX DE Sequence encoding pregastric lipase.
XX KW Precursor polypeptide; secretion vector; enzyme; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT sig_peptide 47..103 /*tag= a
FT mat_peptide 104..1243 /*tag= b
XX W08603778-A.
XX PN 03-JUL-1986.
XX PD 23-DEC-1985; 85WO-GB00599.
XX PF 21-DEC-1984; 84GB-0032483.
XX PR 23-DEC-1985; 85WO-GB00599.
XX PR 01-JAN-1986; 86GB-0019568.
XX PA (BREW-) BREWING RES FOUND.
XX PA (TUBB/) TUBB R S.
XX PI Tubb RS;
XX WPI; 1986-182910/28.
XX P-PSDB; R4P60724.

XX New precursor polypeptide of defined sequence - and corresp. DNA
 PT used to transform hosts for prodn. of the polypeptide
 XX Example; Fig 5; 60pp; English.

XX The inventors claim a DNA sequence encoding AAP60723 linked to a
 CC promoter upstream and a gene for a polypeptide downstream.
 CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,
 CC gastric lipase and the mammalian lymphokine, interferon-alpha2.
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

Query Match 37.0%; Score 470; DB 7; Length 1367;

Best Local Similarity 64.2%; Pred. No. 3.5e-139;

Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtggaccagagcattcatgaattattgtaaaatcaatcaacatcaagctatccct 178
 DB 123 ctggagccctgaagtactatgaacattagtcagatgattactatttggggataccaa 182
 QY 179 gtgaggaatagaagtcgaactgaagatgggtatattcttcttgaacaggattccctc 238
 DB 183 atgaagaatagaagtgtgactgaagatggttatattcttgaagtcataagaattccct 242
 QY 239 gaggcctagtccaaccttaagaagacaggttccaggcctgtgtgttactcagcatgccc 298
 DB 243 atgggaagaataattcagggaatacagccagagacctgtgtgttttggcagcatggtt 302
 QY 299 tagttggaggtcgtcagcactgatttccaaacctgcccacaaatagctgggttcattc 358
 DB 303 tgcctgcatcgcacaaactggatttccaaacctgcgaacagccttcccttcattc 362
 QY 359 tggcagatgctgttttgcgtggaagggaacagcaggggaacccctgtgtctcgaa 418
 DB 363 tggcagatgctgtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 422
 QY 419 aacacaagacacttccatagaccagatgagttctgtgtgtgtgtgtgtgtgtgtgtgt 478
 DB 423 gaaactgttactattcaccagatcagttgaattctgtgtgtgtgtgtgtgtgtgtgtgt 482
 QY 479 ctagggttgaccttctcagtgataaacttatttttgcagaaacggccagagaaaga 538
 DB 483 ctataatgaccttccagcccaactcagcttctcattgttaaagaaactggacagagcag 542
 QY 539 tctattatgctggtcattcagcagccaccatcctgtgtgtgtgtgtgtgtgtgtgtgtgt 598
 DB 543 tacactatgttggccattccagggcagccaccattgtgtgtgtgtgtgtgtgtgtgtgtgt 602
 QY 599 cagagctggctcagaaaaatcaaaatgtatttggcttttagccccaatagccactgttaagc 658
 DB 603 ccagctggctcaaaagaatacaaaacctctatgctctagctcctgtgtgtgtgtgtgtgt 662
 QY 659 atgcaaaagcccgccggaacaaattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 718
 DB 663 atacaaaaagcttataaacaacttagattgttctcctcaactcctctcagtttatat 722
 QY 719 ttggcaaaaaagaatttctgtatcagaccagatttctcagacaactgtgtatttaccctt 778
 DB 723 ttgttgacaaataatttaccacacaaacttctttagatcaatttctgtctactgaagtgt 782
 QY 779 gtggccaggtgatttctgacagattttagtaataatcatgttacttctgtgtgtgtgtgt 838
 DB 783 gtctccgtgagatgtgaattccttcttgcagcaatgccttatttataattgtgtgtgtgt 842
 QY 839 acacacaataatgaatgacgagcagcaagatgtatgtgtgtgtgtgtgtgtgtgtgtgtgt 898
 DB 843 acagtaagaacttcaacagagtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 902
 QY 899 catctgtgcaaaaatttctacacgagccagcaggtgaattctgtgtgaactccgggcat 958
 DB 903 ctctctgtaaaaaacatgttccatttggaccagcaggtgtgttaagctcttgggaaattccaa 962

QY 959 ttgactgggggagtgagacccaaaaatctggaaaaatgcaatcagccaaactcctctgaagt 1018
 DB 963 atgactgggggaagccagttcagaatgagtgactatgatcagtcaccaactcctact 1022
 QY 1019 acagagtcagagatatgacgtccctacagcaaatgtgacagaggtcagagactgcttt 1078
 DB 1023 acaatgtgacagccatgaatgaccaattgacagtggtggaacggttggcaaggacctgttgg 1082
 QY 1079 caaatccagaagcgtgaaaaatgctctctctgagtgaccacacccctcatctaccataaga 1138
 DB 1083 ctgacccccaagatgttggccttttcttccaaaactcccaactttattaccacaagg 1142
 QY 1139 atattcctgaatgggtcactcagtgatttctctctgtgtgtgtgtgtgtgtgtgtgtgt 1198
 DB 1143 agattcttttacaatcacttggactttatctgttgcaatgtgtccctcctcaagaagttt 1202
 QY 1199 acaatgaaatcatcctctgat 1220
 DB 1203 acaatgacattgtttctatgat 1224

RESULT 5

AAN60566

ID AAN60566 standard; DNA; 1367 BP.

XX AAN60566:

XX AC AC

XX XX 22-AUG-1991 (first entry)

XX DT 22-AUG-1991 (first entry)

XX DE Sequence encoding human pregastric lipase.

XX XX Cystic fibrosis therapy; enzyme; lipase deficiency; ss.

XX KW Homo sapiens.

XX OS Homo sapiens.

XX FH Key

XX CDS Location/Qualifiers

FT 47..103

FT /*tag= a

FT 104..1243

FT mat_peptide

FT /*tag= b

XX WO8601533-A.

XX 13-MAR-1986.

XX 15-AUG-1985; 85WO-GB00364.

XX 21-AUG-1984; 84GB-0021210.

XX 15-AUG-1985; 85WO-GB00364.

XX 01-JAN-1986; 86GB-0008897.

XX (CELL-) CELTECH LTD.

XX (LOWE/) LOWE P A.

XX Lowe PA;

XX WPI; 1986-081634/12.

XX P-PSDB; AAP60658.

XX DR New gastric lipase protein, esp. of human origin - for treating

XX PT lipase deficiency, and DNA sequences coding for it

XX PS Disclosure; Fig 3; 39pp; English.

XX CC The inventors claim a pregastric lipase protein and a gene encoding

XX CC it. Gastric lipase protein is useful for oral administration to

XX CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.

XX SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

Query Match

37.0%; Score 470; DB 7; Length 1367;

Best Local Similarity 64.2%; Pred. No. 3.5e-139; Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;			
QY	119	ctgtgacccagagcattcatgaatattagtgaaatcatcccaacatcaaggtatccct	178
DB	123	ctggaagccctgaagtgaactatgaacatttagtcagatgattactatttgggataccaa	182
QY	179	gtgagaataatgaatgcgaatgaatgaatgggtatattctttctgttaacagattccct	238
DB	183	atgaagaataatgaattgtgactgaagatggttatattcttgaagtcacatagaattccct	242
QY	239	gagccctagtgcaactaagaacacaggttccaggctgtggtgttaactgcagcatggcc	298
DB	243	atgggaagaataattcagggaatcacaggccagagagatgtgtgttttgcagcatgggt	302
QY	299	tagttgaggtgctagcaactgatttccaaacctgcgaacatgctgggcttcatc	358
DB	303	tgcttgatcagccacaaactgatttccaaacctgcgaacacagccttgcttcatc	362
QY	359	tggcagatgctgtttttgacgtgtgagtgagggaacagcggggaacgctgtgtctcgaa	418
DB	363	tggcagatgctgtgtatgt	422
QY	419	aacacaaagacactctcatcagacacagatgattgtgtgtgtgtgtgtgtgtgtgtgt	478
DB	423	gaactgtactattcaccagattcagttgaaattctgtgtgtgtgtgtgtgtgtgtgtgt	482
QY	479	ctaggttgacctctctgcagtgatgaataactttatttgcagaacagggccaggaaga	538
DB	483	ctaaatgaccttccagccacaaactgacttcttgaagaacaaactggacagaagcagc	542
QY	539	tctattgtggtctattcagcgggacccaccatgggtttattgcattttccaccatgc	598
DB	543	tacactatgtggcattccagggcaccaccattgtgtttattgtcttttccaccatc	602
QY	599	cagagctgctcagaaatacaaatgtatttctgtttagcaccatagcactgttaagc	658
DB	603	ccagctggctcaaaagatacaaacctctctgtctgtgtgtgtgtgtgtgtgtgtgtgt	662
QY	659	atgcaaaagcccgagcaaaattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	718
DB	663	atacaaaagccctataaacaacattagattgttctcctcaatcccttcaagtattat	722
QY	719	tggcaaaaagaatttctgtatcagacagatttctcagacaaactgttatttaccctt	778
DB	723	tgtgtgcaaaatatttccaccacacacttcttctgtgtgtgtgtgtgtgtgtgtgtgt	782
QY	779	gtggccaggtgcttctgtatcagatttgtgtatatactgttacttctgtgtgtgtgtgt	838
DB	783	gtctccgtgagatgt	842
QY	839	acacaaacaaatgaacatgagccgagcaagtgatatgtgtgtgtgtgtgtgtgtgtgt	898
DB	843	acagtaagaactttaacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	902
QY	899	catctgtcaaaaatttctacactgagccagcagtgaaattctgtgtgtgtgtgtgtgtgt	958
DB	903	cttctgtcaaaaactgttccattgagccagcagtggtgtgtgtgtgtgtgtgtgtgtgt	962
QY	959	tgtactggggagtgagacaaaatactggaataatgcaatcagcacaactcctgtgaagt	1018
DB	963	atgactgggaagccaggttccagaggtgactatgactatgactatgactatgactatgact	1022
QY	1019	acagagtcagagatgacagcttccatcagcaatgtgtgtgtgtgtgtgtgtgtgtgtgt	1078
DB	1023	acaatgtgacagccatgaattgtacaaattgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1082
QY	1079	caaatccagagagcgtgaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1138
DB	1083	ctgaccccaagatgt	1142
QY	1139	atattctgaaatgggtcagctgatttctatctgtgtgtgtgtgtgtgtgtgtgtgtgt	1198
DB			

Db	1143	agattcctttttacaaatcacttgaccttatctgtggcaatgagtgccctcaagaagttt	1202
QY	1199	acaatgaaatcactcactctgat	1220
DB	1203	acaatgacattgtttctatgat	1224
RESULT 6			
ID	AAT58916		
XX	AAT58916 standard; DNA; 1367 BP.		
AC	AAT58916:		
DT	19-AUG-1997 (first entry)		
DE	Human gastric lipase coding sequence.		
KW	Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;		
KW	food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;		
KW	hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.		
OS	Homo sapiens.		
PH	Key	Location/Qualifiers	
FT	CDS	47..1243	
FT		/tag= a	
FT	sig_peptide	47..103	
FT		/tag= b	
FT	mat_peptide	104..1240	
FT		/tag= c	
XX	W09633277-A2.		
PN	XX		
PD	24-OCT-1996.		
XX	XX		
PF	19-APR-1996; 96WO-FR00606.		
XX	XX		
PR	20-APR-1995; 95FR-0004754.		
XX	XX		
PA	(BIOC-) EIOCEM SA.		
PA	(LJOU) INST RECH JOUVEINAL.		
XX	XX		
PI	Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;		
PI	Merot B;		
DR	WPI; 1996-485783/48.		
DR	P-PSDB; AAW09383.		
XX	XX		
PT	Recombinant human or canine pre-duodenal lipase prodn. in transgenic		
PT	plants - useful for facilitating absorption of fat, as bio-catalysts		
PT	and for prodn. of bio-fuel.		
XX	XX		
PS	Claim 3; Fig 4; 130pp; French.		
XX	XX		
CC	This is the nucleotide sequence encoding the human pre-duodenal (i.e.		
CC	gastric) lipase enzyme. The sequence can be used to generate transgenic		
CC	plants producing recombinant lipase in an enzymatically active form.		
CC	Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of		
CC	the mature protein, respectively) can be deleted to form the derivatives		
CC	designated delta-4 or delta-54 respectively. Plants, or their extracts,		
CC	expressing the lipases or the truncated derivatives, can be used:		
CC	(a) as pharmaceuticals or food to facilitate absorption of fat, either		
CC	in healthy subjects or in patients with inadequate levels of		
CC	gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine		
CC	pancreatic insufficiency, in the elderly or in patients undergoing medical		
CC	treatment which alters fat adsorption; (b) for performing industrial or		
CC	agricultural reactions, e.g. in processing of fats or in the dairy		
CC	industry, for hydrolysis or transesterification reactions, etc., where		
CC	the plant material may provide both enzyme and substrate. The transgenic		
CC	plants can also be used for biofuel production.		
XX	XX		
SQ	Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;		

Query Match 37.08; Score 470; DB 17; Length 1367;
Best Local Similarity 64.2%; Pred. No. 3.5e-139;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtggaccagagcattcatgaataattagtgaaatcatccaacatcaaggtatccct 178
DB 123 ctggaaacctgaagtactatgaacattagtcagatgaatactattggggataccaa 182

QY 179 gtgaggaatagagtcgcaactgaagatgggtatattcccttctgttaacagagattctc 238
DB 183 atgaagaatagagtgtagtgagtgagtgatgttatattctgaagtcgaatagaattcctt 242

QY 239 gaggcctagtgcacacctgaagaagaggttccaggccctggtggtgttactcagcagatgcc 298
DB 243 atgggaagaaattccgggaatacagagccagagacctgtgtgttttgcagcatggtt 302

QY 299 tagtggaggtgctagcaactggaattccaaacctgcccacaaatagcctgggttccattc 358
DB 303 tgcctgcatcagccacaaactggattccaaacctgcgaacacagccttgccttcattc 362

QY 359 tggcagatgctggttttgagctgtggatggggaacagcagggaacgctgtctcgaa 418
DB 363 tggcagatgctggttttgagctgtggatggggaacagcagggaacacacctgggccaaga 422

QY 419 aacacagacactctccatagaccacagatgagttctgtggtttcagttatgatgagatgg 478
DB 423 gaaactgtactattccagattcagttgaaattcgtgttccagcttgcagttgagaaatgg 482

QY 479 ctagggttgccttccctgcagtgataaaactttattttgcagaaaaacggccaggaaga 538
DB 483 ctaaatagacctccacggccacacatgcacttcattgtaaaagaaactggacagaagcagc 542

QY 539 tctattatgtcgttctacagggccaccacattggttctgttccaccatagccactgttaagc 658
DB 543 tacactatgtggccattccagggccaccacattggttctgttccaccatagccactgttaagc 662

QY 659 atgcacaaagcccgagacaaattttgtgtgctgcagatatgatatacaaggatgtgt 718
DB 663 atacaaagcccttataacaaacttagattgttctcactccctcttccaaagttatat 722

QY 719 ttggcaaaaagaattctgtatcagaccagattctcagacacactgttattacattt 778
DB 723 ttggtgacaaatattctaccacacacacttcttgatcaatttctgttactgtaagtg 782

QY 779 gtggccaggtgattctctgacagatttctgtatgaataatctgttactctgtgggtgattca 838
DB 783 gctcccgtagatgctgaattctctctgcagcaatgcttatttataatttgggatttg 842

QY 839 acacacaaatatagaacatagccagcgaagtgatgtatgctgccacacacttctgctgaa 898
DB 843 acagtgaagacttcaacacagctgctggtggtgtatctatcacataatccagcaggaa 902

QY 899 catctgtcaaaaattctacactggagccagggcagtgaaattctgtggaactccggcat 958
DB 903 ctctgttcaaaaacttccattggaccagggctgtaagtgctgggaattccaaagctt 962

QY 959 ttgactggggagtgagacaaataatctgaaaaatgcaatcagccaaactctgtgaaggt 1018
DB 963 atgactgggaagcccggttcagaataggtgactatgatcagatgccaaactccctact 1022

QY 1019 acagagtcagagatatgacgttccctacagcaatgtggacagaggtcaggaactggcttt 1078
DB 1023 acaatggacagcagcatgaagtacaaattgcagttggaacggtggcaaggactgttg 1082

QY 1079 caaatccagagagcgtgaaaaatgctctctgtgagtgacacacactctatccataaga 1138
DB 1083 ctgaccccaagatggttggcctttcttccaaaactccccaactctctatttaccacaggt 1142

QY 1139 atattcctgaatgggtcagctgatttcattcattcctgggttggatgctcctcaccgtatgt 1198
DB 1143 agattccttttacaatcacttgactttattcttctgggcaatggatgcccctcaagaagttt 1202

QY 1199 acaatgaatacatccatctgat 1220
DB 1203 acaatgacattgtttctatgat 1224

RESULT 7
AAN50385
ID AAN50385 standard; DNA; 1336 BP.
XX
AC AAN50385:
XX
DT 17-JAN-1992 (first entry)
XX
DE Rat prelingual lipase gene.
XX
KW Prelingual lipase; enzyme; EC-3.1.1.3; ss.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
CDS 7..1194
FT /*tag= a
XX
PN GB214233"-A.
XX
PD 16-JAN-1985.
XX
PF 29-JUN-1984; 84GB-0016581.
XX
PR 01-JUL-1983; 83GB-0017989.
PR 05-SEP-1983; 83GB-0023759.
XX
PA (CELL-) CELTECH LTD.
XX
PI Carey NH, Williamson R;
XX
DR WPI; 1985-014450/03.
P-PSDB; AAP50322.
XX
PT New lingual lipase protein for treatment of lipase deficiency - also
new pre-lingual lipase protein and related products
XX
PS Disclosure; Fig 5; 15pp; English.
XX
CC This sequence may be expressed in a transformant host organism using
a vector plasmid in order to produce a prelingual lipase protein.
CC The expressed protein may be used for the treatment of lipase
deficiency.
XX
SQ Sequence 1336 BP; 409 A; 285 C; 263 G; 379 T; 0 other;

Query Match 36.6%; Score 464.4; DB 6; Length 1336;
Best Local Similarity 63.4%; Pred. No. 2.1e-137;
Matches 711; Conservative 0; Mismatches 411; Indels 0; Gaps 0;

QY 119 ctgtggaccagagcattcatgaataattagtgaaatcatccaacatcaaggtatccct 178
DB 80 ctggaaacctgaagcaaatgaataattagtcagatgaataacttactggtggatccat 139

QY 179 gtgaggaatagagtcgcaactgaagatgggtatattccttctgttaacagagattctc 238
DB 140 gtcaagaatagagttgttactgaagatggtctacattctgggtgtctacagaattctc 199

QY 239 gaggcctagtgcacacctgaagaagcaggttccagcctggtgttactcagcagatgcc 298
DB 200 atgggaagaataattctgaaataataggaagacctgtggtgtatttgcagcatggt 259

Qy 425 agaacctctccatagacacagatgagttcttgggtttccagttatgatgagatgaggtt 484
 Db 326 tgtactactccgactccgctgaaattcttgggtttccagttatgatgagatgaggtt 385
 Qy 485 ttgacctctctcagtgatgaacttttttttcagaaacggccaggaagaagatctatt 544
 Db 386 atgacctctccgacacactctctctctctctctctctctctctctctctctctct 445
 Qy 545 atgtcgggtattacacagggaccacacattgttattgtatttttccaccatgccagagc 604
 Db 446 acgttgccattccaggccacacacattgttattgtatttttccaccatgccagagc 505
 Qy 605 tggctcagaaatacaaaatgtattttgttttagcaccacacacacacacacacacac 664
 Db 506 tggcgaacaggtacaaacactctctctctctctctctctctctctctctctctctct 565
 Qy 665 aaagcccgccgacacaaattttttgtctgcagatatgatgataaaggatgtttggca 724
 Db 566 aaacctgtttaacaaac 625
 Qy 725 aaaaagaatttctgtatcagaccagattttctcagacaaactgtttatttaccctttggcc 784
 Db 626 acaaaattctaccac 685
 Qy 785 aggtgattctgacagattgtgtaataatcatttcttctctctctctctctctctctct 844
 Db 686 gcgagacggtggtatctctctcagacacacacacacacacacacacacacacacac 745
 Qy 845 acaatgaatcagacgagcaagtgatatgtctgcccacacacacacacacacacacac 904
 Db 746 tgaacttgaaac 805
 Qy 905 tgcgaatattctacactggcagcagtggaattctgtggaactccggcatttgact 964
 Db 806 ttcaagacgtctccactggtccacagctgttaagtctgggaagttccaaactttgact 865
 Qy 965 ggggagtgagacacaaataatggaaaaatgcataatcagcacaactctctgtagggatagag 1024
 Db 866 ggggagccaggttcagac 925
 Qy 1025 tcagagatatcagctccctcagacacacacacacacacacacacacacacacacacac 1084
 Db 926 tgcagac 985
 Qy 1085 cagaagacgtgaaaatgctctctgagtgagcacaacacacacacacacacacacacac 1144
 Db 986 ctacagatgtgacctttgtcttccaaagctcccaactctcatttaccacaggaagattc 1045
 Qy 1145 ctgaatgggctacagtgatttctctctctctctctctctctctctctctctctctctct 1204
 Db 1046 ctcttacaatcacttgactttattctggccatggtgcccctcaagcgggtttacaatg 1105
 Qy 1205 aaatcatcatctgatg 1221
 Db 1106 aaattgtttccatgatg 1122

RESULT 9
 AAT58915
 ID AAT58915 standard; cDNA to mRNA; 1528 BP.
 XX
 AC AAT58915;
 XX
 DT 19-AUG-1997 (first entry)
 XX
 DE Dog gastric lipase coding sequence.
 XX
 KW Dog; duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; ss.
 OS Canis familiaris.

FH Key Location/Qualifiers
 FT CDS 1..1140
 FT /*tag- a
 FT /product= gastric lipase
 XX
 PN W09633277-A2.
 XX
 PD 24-OCT-1996.
 XX
 PF 19-APR-1996; 96WO-FR00606.
 XX
 PR 20-APR-1995; 95FR-0004754.
 XX
 PA (BIOC-) BIOCEM SA.
 PA (LJOU) INST RECH JOUVEINAL.
 XX
 PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 PI Merot B;
 XX
 DR WPI; 1996-485783/48.
 DR P-PSDB; AAM09382.
 XX
 PT Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 XX
 PS Claim 2; Fig 1; 130pp; French.
 XX
 CC This is the nucleotide sequence encoding the dog pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 12 bases of this sequence (encoding amino acids
 CC 1-4 of the protein encoded by this sequence) or bases 1-162 of this
 CC sequence (encoding amino acids 1-54) can be deleted to form the
 CC derivatives designated delta-4 or delta-54 respectively. Plants, or
 CC their extracts, expressing the lipases or the truncated derivatives, can
 CC be used: (a) as pharmaceuticals or food to facilitate absorption of fat,
 CC either in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX
 SQ Sequence 1528 BP; 392 A; 387 C; 329 G; 420 T; 0 other;

Query Match 35.0%; Score 444.2; DB 17; Length 1528;
 Best Local Similarity 62.8%; Pred. No. 6.1e-131;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 125 acccagaagcattcatgaattagtgaaatcattccaaatcattcaagctatccctgtgag 184
 Db 26 accctgaagtgcacatgaataatagtcagatgatacctactggggataccacagctgag 85
 Qy 185 aatgtgaagtgcgaactgaagtgggtatattcttttgaacaggattctctcaggcc 244
 Db 86 aatgtgaagtgtgacccaagacggttatattctctgggacgacagaattctcttgga 145
 Qy 245 tagtgcaaccttaagaagacaggttccaggcctgtgttactcagcatggcctagtgtg 304
 Db 146 ggauaaattcagagaataatagccggagacctgtgtgcattttgcacacggtttgctcg 205
 Qy 305 gaggtgtcagcaactggatttccaaacctgcccacaaatagcctggcttcattctggcag 364
 Db 206 catcagccacaaactggatctctccaaacctgcccacaaaggcctggccttcactggcgg 265
 Qy 365 atgtcgggttttgacgtgtggatggggaacagcaggggaacacccctggttctcgaaacaca 424
 Db 266 acgcccgggttacgacgtgtggctgggggaacagcaggggaacacacctggggcaggaggaatc 325

Oy 583 gcatttc 590
| || ||
Db 754 tccttc 761

Search completed: August 24, 2001, 22:42:12
Job time: 7433 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 21:22:40 : Search time 2518.61 Seconds
(without alignments)
4762.814 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgcaag.....gacgggtgagccgtattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:
12: gb_est12:
13: gb_est13:
14: gb_est14:
15: gb_est15:
16: gb_est16:
17: gb_est17:
18: gb_est18:
19: gb_est19:
20: gb_est20:
21: gb_est21:
22: gb_est22:
23: gb_est23:
24: gb_est24:
25: gb_est25:
26: gb_est26:
27: gb_est27:
28: gb_est28:
29: gb_est29:
30: gb_est30:
31: gb_est31:
32: gb_est32:
33: gb_est33:
34: gb_est34:
35: gb_est35:
36: gb_est36:
37: gb_est37:
38: gb_est38:
39: gb_est39:
40: gb_est40:
41: gb_est41:
42: gb_est42:
43: gb_est43:
44: gb_est44:
45: gb_est45:
46: gb_est46:
47: gb_est47:

44: em_esthum10:
45: em_esthum11:
46: em_esthum12:
47: em_esthum13:
48: em_esthum14:
49: em_esthum15:
50: em_esthum16:
51: em_esthum17:
52: em_esthum18:
53: em_esthum19:
54: em_esthum20:
55: em_esthum21:
56: em_esthum22:
57: em_esthum23:
58: em_esthum24:
59: em_esthum25:
60: em_esthum26:
61: em_esthum27:
62: em_esthum28:
63: em_estin1:
64: em_estin2:
65: em_estin3:
66: em_estin4:
67: em_estin5:
68: em_estomi:
69: em_estov1:
70: em_estov2:
71: em_estov3:
72: em_estpl1:
73: em_estpl2:
74: em_estpl3:
75: em_estpl4:
76: em_estpl5:
77: em_estpl6:
78: em_estpl7:
79: em_estpl8:
80: em_estpl9:
81: em_estpl10:
82: em_estro1:
83: em_estro2:
84: em_estro3:
85: em_estro4:
86: em_estro5:
87: em_estro6:
88: em_estro7:
89: em_estro8:
90: em_estro9:
91: em_estro10:
92: em_estro11:
93: em_estro12:
94: em_estro13:
95: em_estro14:
96: em_estro15:
97: em_estro16:
98: em_estro17:
99: em_estro18:
100: em_estro19:
101: em_estro20:
102: gb_est25:
103: gb_est26:
104: gb_est27:
105: gb_est28:
106: gb_est29:
107: gb_est30:
108: gb_est31:
109: gb_est32:
110: gb_est41:
111: gb_est42:
112: gb_est43:
113: gb_est44:
114: gb_est45:
115: gb_est46:
116: gb_est47:

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	618.2	48.7	2927	192 AK019504	Mus muscu
2	505	39.8	954	144 BF135102	BF135102 601780142
3	497.6	39.2	1079	145 BF141567	BF141567 601788654
4	471.8	37.2	669	145 BF143546	BF143546 601790896
5	461.8	36.4	1350	192 AK010093	Mus muscu
6	460.2	36.3	1347	192 AK010116	Mus muscu
7	460.2	36.3	1349	192 AK009300	Mus muscu
8	460.2	36.3	1350	192 AK010139	Mus muscu
9	458.6	36.1	1340	192 AK010124	Mus muscu
10	458.6	36.1	1343	192 AK009479	Mus muscu
11	458.6	36.1	1345	192 AK009413	Mus muscu
12	458.6	36.1	1345	192 AK009473	Mus muscu
13	458.6	36.1	1346	192 AK009359	Mus muscu
14	458.6	36.1	1346	192 AK009523	Mus muscu
15	458.6	36.1	1346	192 AK009729	Mus muscu
16	458.6	36.1	1348	192 AK010058	Mus muscu
17	458.6	36.1	1349	192 AK009474	Mus muscu
18	458.6	36.1	1349	192 AK009546	Mus muscu
19	458.6	36.1	1349	192 AK009571	Mus muscu
20	458.6	36.1	1349	192 AK009573	Mus muscu
21	458.6	36.1	1349	192 AK010148	Mus muscu
22	458.6	36.1	1350	192 AK009459	Mus muscu
23	458.6	36.1	1350	192 AK009525	Mus muscu
24	458.6	36.1	1350	192 AK009773	Mus muscu
25	458.6	36.1	1350	192 AK010061	Mus muscu
26	458.6	36.1	1350	192 AK010125	Mus muscu
27	458.6	36.1	1350	192 AK010203	Mus muscu
28	458.6	36.1	1351	192 AK009428	Mus muscu
29	458.6	36.1	1352	192 AK009431	Mus muscu
30	458.6	36.1	1356	192 AK010019	Mus muscu
31	458.6	36.1	1356	192 AK010035	Mus muscu
32	457.6	36.1	1346	192 AK009437	Mus muscu
33	457.6	36.1	1347	192 AK010103	Mus muscu
34	457.6	36.1	1349	192 AK010026	Mus muscu
35	457	36.0	1346	192 AK009560	Mus muscu
36	457	36.0	1347	192 AK010236	Mus muscu
37	457	36.0	1348	192 AK010231	Mus muscu
38	457	36.0	1349	192 AK009875	Mus muscu
39	457	36.0	1350	192 AK009537	Mus muscu
40	456	35.9	1349	192 AK010106	Mus muscu
41	455.4	35.9	1345	192 AK009544	Mus muscu
42	455.4	35.9	1349	192 AK009990	Mus muscu
43	453.8	35.8	1346	192 AK010173	Mus muscu
44	449.8	35.4	1407	192 AK009923	Mus muscu
45	448.2	35.3	1350	192 AK010123	Mus muscu

ALIGNMENTS

RESULT	1	
AKO19504		
LOCUS		
DEFINITION		AKO19504 2927 bp mRNA HTC 08-FEB-2001 Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23, full insert sequence.
ACCESSION		AKO19504
VERSION		AKO19504.1 GI:12859754
KEYWORDS		CAP trapper.
SOURCE		Mus musculus (strain:C57Bl/6J) 0 day neonate skin cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:4632427C23.
ORGANISM		Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE		1 (sites)
AUTHORS		Carninci,P. and Hayashizaki,Y.
TITLE		High-efficiency full-length cDNA cloning

JOURNAL REFERENCE
AUTHORS

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

JOURNAL MEDLINE REFERENCE
AUTHORS

3 (sites)

Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Hirachi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Hkuda,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kawahagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913

JOURNAL MEDLINE REFERENCE
AUTHORS

4 (sites)

The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 2927)

Aéchi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Hanakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Scabey,X., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,Y., Tejima,Y., Toya,Y., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN). Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

JOURNAL REFERENCE
COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCTCGAGTTAATTAAATTCACCCCCCCCCC 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATTAAATTCACCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES source

Location/Qualifiers

1..2927

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="MGI:1912763"

/db_xref="MGI:1926003"

/clone="4632427C23"

/tissue_type="skin"

/clone_lib="RIKEN full-length enriched mouse cDNA library"


```

Db 241 TTTTATACAGACTAGATTTTTCAGGAGCTTTTATTTTACC-TTGTGGCCAGATGATT 299
Qy 793 ctgtatcagattttagtaatatcatgttacttctgtgtgtgattcaacacccaataatg 852
Db 300 CTCGACCAAAATCTGCAGCAACATCATCTTACTCTCTGGGAGGTTTAAACACAAACAAATATG 359
Qy 853 aacatgagccagagcaagtgtatgtctgcccacactcttctgtgaaacatctgtgcaaaat 912
Db 360 AACATGAGCAGAGCAAAATGTGTGTGCCCCATACACTGTGGAAGCTCTGTGCAGAAC 419
Qy 913 attcactagagccagcgagtaattctgtgtaacctccgggcatttgaactgagggaggt 972
Db 420 ATTCTCACTGGAGCCAGCGCGTGAATCTGGGAACCTCTGTGCTTTGACTGGGGAAGT 479
Qy 973 gagacccaaaaatctggaaaaatgcaatcagccaaactcctgtgaaggtcagagtcagagat 1032
Db 480 GAGACCAAAAATCAGGAGAAATGTAATCAGCCAACCTCTTAAAGGTACAAAGTTCGAGAT 539
Qy 1033 atgacggtccctacagaatgtgacagggaggtcaggagctggttcaaatccagaagac 1092
Db 540 ATGATGGTCCCAACAGCAATGTGACTGGAGGTCAAGACTGG-TTCAAAATCCAGATGAT 598
Qy 1093 gtgaaaatgctctctctgaggtgaccaaactcctacataagaatattcctggaatgg 1152
Db 599 GTGAACAATTTACTTCTGAAGTACCAAACTCATCTACCAAGACATCTCTGAGGTG 658
Qy 1153 gctcaatggtatttcattctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1193
Db 659 GGTATGTGGA-TTCATCTGGGGGGTGGATGCCCTCAGCG 698

RESULT 3
BF141567 1079 bp mRNA EST 24-OCT-2000
LOCUS 601788654F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5',
DEFINITION mRNA sequence.
ACCESSION BF141567
VERSION BF141567.1 GI:10980607
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1079)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lni.gov
Plate: L1AM9263 row: 1 column: 14
High quality sequence stop: 663.
Location/Qualifiers
1..1079
/organism="Mus musculus"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:4016413"
/clone.lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; transgenic model WNT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
312 a 267 c 263 g 237 t

```

ORIGIN

Query Match 39.28; Score 497.6; DB 145; Length 1079;
 Best Local Similarity 84.18; Pred. No. 3.4e-140;
 Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;

```

Qy 256 aagsagacagattccagggcctgtgttactgcagcatggcctagttgggaggtgctagc 315
Db 20 AAGAAGGAAGGATCCAGGCCAGTGGTGTACTGCAACATGTCTCTTTGGGGATGCTAGC 79
Qy 316 aactggatttccaaacctgcccacaacatagcctgggtcttcattctggtgcagatgctgtttt 375
Db 80 AACTTGGATTTTCCAACTGCCCCAACACAGCCCTGGGTTTTTATTTCTGCGAGATGCAAGTTT 139
Qy 376 gacgtgtggtggtggaacacagcaggggaaagccc-tggctctgaaacacacagacactctc 434
Db 140 GATCTGTGGATGGGAACACAGCAGAGGAANACACCCTGGTGTCTCGGAAGACACAGCCCTCTC 199
Qy 435 catagaccaagatgagttctgggttttcagttatgatgagatggctaggttttgaccttcc 494
Db 200 CATAGATCAAGATCAGTCTGGGCTTTCAGTTATGATGAATGGCTAGGTTTGACCTTCC 259
Qy 495 tgcagtgataaaacttattttgcgaaacacggccagagaaagatctctattatgctggeta 554
Db 260 AGCTGTGATAAACTTTATCTTACAGAAACGGCCAGAAAGGTCTATTATGCGGCTA 319
Qy 555 ttcacagggcacacacacatgggctttattgcatcttccaccatgccagagctggctcagaa 614
Db 320 CTCACAGGGCACCACTATGGCTTTATTTGCAATTTTCCACAATGCCAGAGCTAGCTCATAA 379
Qy 615 aatacaaatgtattttcttttagcaccatagccactgtttaagcatgcaaaaagcccg 674
Db 380 AATCAAAATGTATTTGGCTTAGCTCTATAGCCACTGTTAATAATGCAAGAGTCTCTGG 439
Qy 675 gaccaatttttctgtgtgccagatatgatcaagggtatgtttggcgaataaagatt 734
Db 440 TACCAAATTTCTGTCTGCCAGATATGATGATCAAGGTATTATTTTGCGACACAGAGTT 499
Qy 735 tctgtatcacagagatttctcagacaaactgtttattacaccttggccaggtgattct 794
Db 500 TTTTATCAGACTAGATTTTTCAGGCAGCTTTTATTATTTACCTTTGTGGCAGATGATTCT 559
Qy 795 tgatcagattttagtaatacatctgttacttctgtgtgtgattcaacacccaataatgaa 854
Db 560 CGACCAATCTGCAGCAACATCATCTTACTCTGGAGGTTTAAACACAAACCATATGAA 619
Qy 855 catgagccagca--agtgtatatgtgccacacactcttctgtggaacatctgtg---caa 909
Db 620 CATGAGCAGACCAACATGTGCTCTGCGCCCATAAACCTGTGGAACGCTCTGTGTGCGGA 679
Qy 910 aatttctacactgagccagg-cagtgaattctggtg-aactccgggcatttgaactggg 967
Db 680 ACCATTCTCCACTGGAGCCAGCGCGGTGAATTTGGGGCACCTTGTGCGCTTTGACTGTG 739
Qy 968 ggaagtgaaccacaaatc 985
Db 740 GGCAAGCGACACCAACCAAGC 757

```

RESULT 4

```

LOCUS BF143546
DEFINITION 601790896F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021870 5',
mRNA sequence.
ACCESSION BF143546
VERSION BF143546.1 GI:10982586
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 669)
24-OCT-2000
EST
Mus musculus cDNA clone IMAGE:4021870 5',
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

BASE COUNT

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1M9277 row: o column: 23
 High quality sequence stop: 659.
 Location/Qualifiers
 1. 669
 /organism="Mus musculus"
 /strain="C2B6H II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4021870"
 /clone_lib="NCI_CGAP_Lu30"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; transgenic model WNT-1, expression driven by
 MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
 dt. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 140 c 165 g 169 t
 ORIGIN
 Query Match 37.2% Score 471.8; DB 145; Length 669;
 Best Local Similarity 85.9% Pred. No. 2e-132;
 Matches 535; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
 Qy 1 atgttgaaacctgtcaagacagtgattgtctccacacagaatggaaatgtggtctctg 60
 Db 48 ATGTCAGAAATCTTGTCAAGAGTGTGGACTGTTTCGCACAGAGTGGAGATGGCTCTG 107
 Qy 61 attctgtggtcgtatgtctccagagaatgtgaatcagatcacatagcgaactaaagct 120
 Db 108 ATTCGTGTAGCATATTACTCCAAAGAAATGTGACTCGGCACATTTGCCACAGAAAGCT 167
 Qy 121 gtgaccacagaagcattcatgaattagtgaaatcatccaaatcaagcgtatccctgt 180
 Db 168 GCGGATCCAGAAAGCATTCATGATTTAGCGAAATCATCAACACAGAGGTTATCCAGT 227
 Qy 181 gaggaatatgaatgcgaactgaagtgggtatattctctcttcttaacaggattctctga 240
 Db 228 GAGGAGTATGAAGTTGCAACAGAGATGGGTACATCTCTTGTGAAACAGAAATCCCTCGG 287
 Qy 241 ggcttagtgcacacaaagacaggttccaggcctgtgtgttactgacgagcattggccta 300
 Db 288 GCACAGACACGGTTAAAGAGGAGGATCCAGGCCAGTGTGTACTGCAACATGGTCTT 347
 Qy 301 gttggagtgtagcaactgatttcccaactgcccacaaatagcctgggcttcaattctg 360
 Db 348 CTGTTGGGATGCTAGCAACTGGATTTCCAACCTGCCCAACACAGCCTGGGTTTATTCTG 407
 Qy 361 gcagatgctggttttgacgtgtgtgattggggaacagagggggaacacgctggtctgaaaa 420
 Db 408 GCAGATCAGG-TTTGATGTGTGGATGGGAACACAGCAGAGAAACACCTGGTCTCGAAG 466
 Qy 421 cacagacactctccatagaccagatgattctgggcttcttgattatgatgatgct 480
 Db 467 CACAGACCCCTCTCCATAGATCAGATGAGTCTCGGCTTTCAGTTAATGAAATGGCT 526
 Qy 481 aggttgacctctctgcagtgataaaactttattttgcagaaaaacggggcaggaagaagatc 540
 Db 527 AGGTTTGACCTTCCAGCTGTGATAAATTTATCTTACAGAAAGCGGCCAGAAAGGTC 586
 Qy 541 tattatgtcggctattcacagggcaccacacatgggctttattgtcattttccacacatgcc 600

Db 587 TATTATGTCGGCTACTACAGGCGCACCATGGCTTTATTCATTTCACAATGCCA 646
 Qy 601 gagctggctcagaataatcaaaat 623
 Db 647 GAGCTAGCTCATACATCCAAAT 669
 RESULT 5
 AK010093 LOCUS
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310067K20, full insert sequence.
 AK010093 VERSION
 KEYWORDS CAP trapper.
 SOURCE Mus musculus
 clone:2310067K20.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 REFERENCE Carlinci, P. and Hayashizaki, Y.
 AUTHORS High-efficiency full-length cDNA cloning
 TITLE Methods Enzymol. 303, 19-44 (1999)
 JOURNAL
 REFERENCE 2 (sites)
 AUTHORS Carlinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 REFERENCE 3 (sites)
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 REFERENCE 4 (sites)
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5 (bases 1 to 1350)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanganaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imatani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sojabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
 cdna library was prepared and sequenced in Mouse Genome


```

/clone="2310069p19"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/notes="putative"
/codon_start=1
/protein_id="BAB26725.1"
/db_xref="GI:12845372"
/translation="MWLLLVTSVLSAFGAGHGLFGKLPKPEANNVNSOMITYWYCP
SEYEVVTEDEGVLGVIRPYCKKSENIGKRPVAVLOHGLIASATNWTNLPNSLA
FILADAGVWLGNSGRNKNYVSPDSVEFRAFSPDEMAYDLPAIDFIVQKT
GEKIHVGHSDGTTIGFAFSTNPALAKIRFYALADIAVTKVTPESFPKLSLIPK
FLKVPFGNKMFPNHYLDQFLGTEVCSRELLDLCSNALFICFGDFKKNLNSRFDV
YLGNPAGTSTQDLFWAQLAAKQVNGSPLOKMLHYNKQTPPYVDVSAINTVPI
AVYNGCHDILADPDQVAMLEPLPLNLLYKELLPYNHLDFIWMADAPQEVYNEIVTM
AED"
BASE COUNT      398 a 289 c 283 g 380 t
ORIGIN

```

```

Query Match      36.3%; Score 460.2; DB 192; Length 1350;
Best Local Similarity 63.7%; Pred. No. 9.2e-129;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 125 acccagaagcattcgaatattagtgaaatcatccaacatcaaggtcctccctgtgag 184
DB 132 ACCCTGAAGCAACATGAATGTAGTCAGATGATACTTACTGGGGATATCCAAGTGAGG 191

QY 185 aatagtaagtcgaactgaagtaggggtatattctcttctgttaacagagattcctcgggcc 244
DB 192 AATATGAAGTGTGTACTGAGATGGCTACATCTCTGGGGGTCTATAGAAATTCCTATGGGA 251

QY 245 tagtgcacctaagaagcaggtccagcctgtgtgttactgcacatggcctaggttg 304
DB 252 AAAAAATTCAGATATCGGCAGAGACCTGTGGCATATTTGCGACATGGTTGATG 311

QY 305 gaggtgttagcactgattcctcaacacgtcccaacataagtcctgggtcctcattctggcag 364
DB 312 CATCAGCCACAACTGCATTAACAACTCTGCCAAACACAGCCTGGCCTTCATCTAGCAG 371

QY 365 atgctgttttgacgtgtgagtgaggggaacagcagggggaacacgcctgggtctcgaacacaa 424
DB 372 ATGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGGAATACATGTGTCGGGAATAATG 431

QY 425 agacactctcatagaccaaagtagtctgtgggtcttcagttatgatgagatggctaggt 484
DB 432 TATACTATTACACAGACTCAGTTGAATTCGGGCTTTTCAGCTTTGATGAATGGCTAAAT 491

QY 485 ttgacctctctgagtgatgaataatttttgcagaaacgggcccaggaagaagatctatt 544
DB 492 ATGACCTTCACGCCACCATAGACTTTCATTGTACAGAAACTGGACAGAGAGATACACT 551

QY 545 atgctggtctattcacagggcaccacacatgggtttatttgattttccaccatgccagagc 604
DB 552 ATGTGTGCTACTTCAGGGCACCACACTATCGGTTTATTGTCTTTCTTAQCAATCCTGCTC 611

QY 605 tggctcagaatacaaatgtatttctttagcaccatagcactgacactgttaagcatgcaa 664
DB 612 TGGCTAAANAAYCAGAGGTTTATGTCATAGCTCCAAATGCTGCTGCTGAGTATACAG 671

QY 665 aaagcccccggaccacaaatttttggctgcagatatgatgatgaagggattgtttggca 724
DB 672 AAGTCCCTTTAAAAAGATTTCACTTATTCCTAAGTTTCTCACAAGGTATATTGTA 731

QY 725 aaaaagaattctgtatcacagcagagattctcagacacactgttatttacctttgtggcc 784
DB 732 ACAAATGTTTCATGCCCCACAACTACTTATGATCAATTTCTTGGTACGGAAGTGTGCTCAC 791

QY 785 aggtgattcttgatcagattgttagtaataatcatgttaactcttggtgggttaccacaa 844
DB 792 GGGAGCTGTAGATCTCTCTGACGACCAACGCTTTATTATCATCTTCTGTGGGATTGACAA 851

```

```

QY 845 acaatatgaacatgagcggcaggaagtgtatctgctccacactcttgcgtggaacatctg 904
DB 852 AAAAATTAAATGTGAGTCGCTTTGTATCTGTATCTAGGCATATATCCAGCAGGAACATCTA 911

QY 905 tgcataatattctacactgagcggcagcagcagtgaaattctgtggaactcggggaacttgact 964
DB 912 CTCAAGACCTTTTCCACTGGGCACAGCTTGTCTAAATCTGGGAAGCTTCAAGCCCTATAACT 971

QY 965 gggggagtgagacacaaatcttgaaaaatgcaactcagcacaactcctgtgaagtcacagag 1024
DB 972 GGGGAAGTCCATTACAGAACTGTACACTACATACATACAGAAACCCCTCCCTACTATGATG 1031

QY 1025 tcagagatatgacgggtccctcacagcaatgtggacagaggtcaggaactgggttccaaatc 1084
DB 1032 TGTGAGCCATGACCGTGCCCAATTCAGCTGTGGAAACGGTGGCCATGACATCCTGGCTGATC 1091

QY 1085 cagaagacgtgaaatgtctctctgtgagtgaccacacctcatctacataagaatatttc 1144
DB 1092 CCCAAGATGTGCAATGCTGCTTCCCAAACTCCCAACCTTTTGTACCATTAAGAGATTC 1151

QY 1145 ctgaatgggtgcacgtgatttcattctgggttgggtggtcctcaccgtatgtacaatg 1204
DB 1152 TTCCCTACATACCTGGACTTTCATCTGGCGATGGATGGCCCTCAAGAGGTTTACAAATG 1211

QY 1205 aatcatccatctgtatg 1221
DB 1212 AGATAGTTACCATGATG 1228

RESULT 9
AKO10124 1340 bp mRNA HTC 08-FEB-2001
LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DEFINITION library, clone:2310069D09, full insert sequence.
ACCESSION AKO10124
VERSION AKO10124.1 GI:12845346
KEYWORDS CAP trapper.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Seiuognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1340)

```


Query Match 36.1%; Score 458.6; DB 192; Length 1345;
Best Local Similarity 63.6%; Pred. NO. 2.8e-128;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 125 accagaagcattcatgaattagtgaaatcatccaacatcaagcgtatccctgtgagg 184
DB 131 ACCCTGAAGCAACATGAATGTTGTCAGATGATAACTACTGCGGGATATCAAGTGAGG 190

QY 185 aatgaagtcgcaactgaagtggtatattctctgttaacaggattctcagagcc 244
DB 191 AATATGAAGTGTCTACTGAAGTGTCTACTGCGGGTCTATAGATTTCTTATGGGA 250

QY 245 tagtgaacctgaagcaggtccagcctgtgtgttactcagcagcagcctagttg 304
DB 251 AGAAAAATTCAGGAATATCGCAAGAGACCTGTGGCATATTTCAGCATGCTTTGATTG 310

QY 305 ggggtgctagcaactgatttcccaacctgcccaacaatagcctggtctcattctggcag 364
DB 311 CATCAGCCACAACTGGATTCAAAATCTGCCAAACACAGCCTGGCCCTCATTTCTAGCAG 370

QY 365 atgtgggtttgacgtgtggaaggaacagcaggggaaacgctgtgctcgaacacaca 424
DB 371 ATGCTGGCTATGATGTGTGGTGGGACACAGTCGAGGGAATACATGGTCCCGGAAATG 430

QY 425 agacactctccatagaccacagatgagttctgggtcttcagttatgatgagatggctaggt 484
DB 431 TATACATATTCACGACACTCAGTTGAATCTGTGGCTTTCAGCTTTCATGAATGGCTAAT 490

QY 485 ttgacctcttcagtgatgaacttattttgcagaaacggcaggaagaaagatctatt 544
DB 491 ATGACCTTCCAGCCACCATAGATTCATTCATGACAGAAACTGGACAGAGATACACT 550

QY 545 atgtcggtattcacagggccaccacatggctttattgcatgttttccaccatgccagagc 604
DB 551 ATGTTGGTCACTCTCAGGCGCACCATATCGGTTTTATTGGCTTTTCTACCAATCTCGTCTC 610

QY 605 tggctcagaataatcaaaatgtattttgttttagcaccatagccactgttaagcatgcaa 664
DB 611 TGGCTTAAAAAATCAAGAGGTTTTATGATCATAGCTCCAGTTGCTACTGTGAAGTATACAG 670

QY 665 aaagccgggacccaaattttgtgtgcagatatgatcatcaagggattgttgca 724
DB 671 AAAGTCCCTTAAAAAGATTTCACTTATTCCTAAGTTTCTCAGGTGATATTGGTA 730

QY 725 aaaaagaattctgtatcagaccagattctcagcaactgttattacctttgtgccc 784
DB 731 ACAAAATGTTATGCCCCACACACTACTTAGATCAATTTCTTGGTACGGAAGTGCTCAC 790

QY 785 aggtgattcttgatcagattgttagtaataatcattcttctggtggtgattcaaccca 844
DB 791 GGGAGCTGTAGATCTCTCTGCAGCAACGCTTTATTATCATCTCTCTGTGGATTGACAGA 850

QY 845 acaataatgaacatgagccgagcagtgatgtatgtgcccaactcttctggaacatctg 904
DB 851 AAAAATTAATGTGTAGTGTGCTTGTATGTGTATAGGGCATAAATCCAGCAGGAACATCTA 910

QY 905 tgcataatattctacactggagccagcagtgaaattctgttggaactcgggcatttgact 964
DB 911 CTCAGACCTTTTCCACTGGGCACAGCTTGTAAATCTGGGAAGCTTCAAGCCCTATAACT 970

QY 965 gggggagtgagaccacaaatctggaaaaatgcgaactcagccaactcttaagggtacagag 1024
DB 971 GGGGAAGTCCATACAGAACATGTTTACACTACATCAAAACGCTCCCTACTATGATG 1030

QY 1025 taagagatatagcgtccctcagacaatgtggacagggaggtcagagactggcttcaaatc 1084
DB 1031 TGTACGCAATACCCCTGCAATTTGAGTGTGGAGCGGTGGCCATGACATCTCGGTGATC 1090

QY 1085 cagaagacgtgaataatgtctctctgtgagtgagcaaacctcattctacacataagatttc 1144
DB 1091 CCCAAGATGTCGCAATGCTGCTTCCCAAACTCCCAACCTTCTGTACCAATGAGGAGATTC 1150

QY 1145 ctgaatgggctcagctggtgatttcattctcctgggttggtgattcctcaccgtatgtacaatg 1204
DB 1151 TTCCCTACAATCACCTGACTTCATCTGGCGATGATGCTCAAGAGGTTTACAATG 1210

QY 1205 aaatcatcattctgatg 1221
DB 1211 AGATAGTTACCATGATG 1227

RESULT 12
AK009473 1345 bp mRNA HTC 08-FEB-2001
LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022H17, full insert sequence.
ACCESSION AK009473
VERSION AK009473.1 GI:12844293
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:2310022H17.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
REFERENCE Carninci,P. and Hayashizaki,Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE Methods Enzymol. 303, 19-44 (1999)
JOURNAL
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1345)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Inotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sugabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gs.riken.go.jp, URL:http://genome-gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for


```
/db_xref="MGI:1898390"
/db_xref="MGI:1914967"
/clone="2310040L03"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN_full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
CDS
     /note="putative"
     /codon_start=1
     /protein_id="BAB26466.1"
     /db_xref="GI:12844703"
     /translation="MWLLVTSVLSAFGCAHGLFGKLPKNPKNMNVNQMITWGY
SEEYEVVDEGIILGVIRPIYKKNSENIGRPVAYLOHGLIASINWITLNPNNSLA
FILADAGVIMLGRSNTWRKNVYSPDSVEFWAFSEDEMAKYDLPATIDFIVQKT
GOEKIHYVHSGQTTIGFATFSTPALAKIKRPYALAPVATVRYTESPKKISLIPK
FLKVIQNKFMFPHNYLDQFLGTEVCSRELLDLICSNALFICGDFDKNNLVSRFDV
VLGHNPACTSTQDILFHWAQLAKSKLOAYNGSPLONNLHYNQKTPPYDYSAITVPI
AVWNGHDLADPDQDVAMLLPKLPNLLYKELPYNHLDFIWMADAPQEVNEYITMM
AED"
BASE COUNT      395 a   288 c   283 g   380 t
ORIGIN
```

```
Query Match      36.1%; Score 458.6; DB 192; Length 1346;
Best Local Similarity 63.6%; Pred. No. 2.8e-128;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

QY 125 acccagaagcattcatgattaggaatcatccaacatcaagcgtatccctgtgag 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 132 ACCCTGAAGCAACATGAATGTTAGTCAGATGATTAATCTTACTGGGGATATCCAAGTGG 191

QY 185 aatataagtcgcaactgaagtggtatctctcttctgttaacaggattccctcaggcc 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 192 AATATGAAGTGTACTGAAGTGGCTACATCTCTGGGGGTCTATAGAATCTCTTATGGA 251

QY 245 tsgtgcacactaagaagacaggttccaggccctgtggtgttactgcagcagtggtcctag 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 252 AGAAAAAATCTGAGAAATATCGGCAAGAGACCTGTGGCATATTTGCAGCATGGTTGATTG 311

QY 305 gagggtctagcaactgatttccaaactgcccacaataagcctggtctcgaagacaca 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 312 CATCAGCCCAAACTGGATTTACAAATCTCGCAAAACACGCTGGCCTTCAATCTAGCAG 371

QY 365 atgctggttttgacgtgtggtggaagacagcaggggaaacgcctggtctcgaagacaca 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 372 ATGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGGAATACATGGTCCCGGAAAAATG 431

QY 425 agacactctccatagaccagaatgagttctgtggtcttccagttatgatgagtggttag 484
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 432 TATACTATTTCACAGACTCAGTTGAATTTCTGGCTTTTACGTTTATGATGAATGGCTAAAT 491

QY 485 ttgacctctctgagtgataaaatttttgcagaaaaacggccaggaagaaagatctatt 544
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 492 ATGACCTTCAGCCACCACATAGACTTCATTTGTACAGAAACTGGACAAGAGAGATACACT 551

QY 545 atgtcggtctattcacagggccacacacctgggctttattgtcattttccaccatgccagagc 604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 552 ATGTTGGTCACTCTCAGGCGACACATATCGGTTTATTGTCCTTTTCTACCAATCTGCTC 611

QY 605 tggctcagaaaaatacaaaatgtattttgttagcaccctagccactgttaagcatgcaa 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 612 TGGCTAAAAAATCAAGAGGTTTATGCAATTTAGCTCCAGTTGCTACTGTGAAGTATACAG 671

QY 665 aaagcccccgggaccacaaattttgtgtcgcagatgatgatcaaggattgtttggca 724
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 672 AAAGTCCCTTTAAAAAGATTTTCACTTATTCCTTAAGTTTCTCTCAAGGTGATATTGGTA 731

QY 725 aaaaaagaattctgtatcagaccagactttctcagacaaactgtttattacctttgtggcc 784
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 732 ACAAAATGTTTCATGCCCACTACTACTAGATCAATTTCTTGGTACGGGAAGTGTGCTCAC 791

QY 785 aggtgattcttgatcagattgttagtaataatcatgttactcttgggtgattcaaaccca 844
```

Search completed: August 24, 2001, 21:23:06
Job time: 5573 sec

```
DB 792 GGGAGCTGCTAGATCTTCTCTGCAGCAGCGTTTATTTCATCTTCTGTGGATTGACAAGA 851
QY 845 acantatgaacatgagccgagcaagtgtatatgtctgcccacactcttctggaacatctg 904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 852 AAAACTTAATGTGAGTCCTCTTGTATCTAGGGCATATCCAGCAGGAACATCTA 911
QY 905 tgcnaaatattctacacatgagccagcagtgaaattctgtgtaactccgggcaatttgact 964
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 912 CTCAGAGACCTTTTCCACTGGGCACAGCTTGCTAAATCTGGGAAGCTTCAAGCCTATAACT 971
QY 965 ggggagtgagaccacaaaatctggaaaaatgcaaatcagccaaactctctgtaaggtacagag 1024
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 972 GGGGAAGTCCATTACAGAACATGTTACACTACAAATCAGAAAAACGCTCCTACTATGATG 1031
QY 1025 tcaagagatatgacgtccctacagcaaatgtgacagagagtcaggactggctttcaaatc 1084
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1032 TGTGAGCCATGACCGTGCCCAATTTGCAGTGTGGAACGGTGGCCATGACATCCTGGCTGATC 1091
QY 1085 cagzagacgtgaaaaatgctctctctgaggtgaccaaactcatctaccataagaatattc 1144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1092 CCCPAGATGTCGCAATGCTGCTTCCCAAACTCCCAACCTTCTGTACCATTAAGGAGATTC 1151
QY 1145 ctgaatgggctcacgtggatttcattctgtgggttttggatgctctcaccctacgtatgacaatg 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1152 TTCCCTACAATCACCTGGACTTTCATCTGGGCGATGGATCGGCTCAAGAGGTTTACAATG 1211
QY 1205 aaatcatccatctctgatg 1221
    ||||| |||||
DB 1212 AGATAGTTACCATGATG 1228
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 20:38:19 ; Search time 211.8 seconds
(without alignments)
6059.636 Million cell updates/sec

Title: US-09-333-159-45
Perfect score: 2044
Sequence: 1 gtcaccacgcgtccggg.....aaaaaaagcgccgc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

- 1: /SID88/gcgdata/geneseq/NA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/NA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/NA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/NA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/NA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/NA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/NA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/NA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/NA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/NA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/NA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/NA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/NA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/NA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/NA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/NA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/NA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/NA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/NA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/NA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/NA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	AAF45131	Human TANGO 294 CD
2	1269	62.1	1269	AAF45132	Human TANGO 294 OR
3	489.2	23.9	1378	AAQ42310	RGL precursor. Or
4	470	23.0	1367	AAN60685	Sequence encoding
5	470	23.0	1367	AAN60566	Sequence encoding
6	470	23.0	1367	AAFS8916	Human gastric lipa
7	464.4	22.7	1336	AAN50385	Rat prelingual lip
8	444.2	21.7	1528	AAQ68388	Canine gastric lip
9	444.2	21.7	1528	AAQ58915	Dog gastric lipase
10	444.2	21.7	1531	AAQ68389	Canine gastric lip
11	286	14.0	699	AAF28679	Human protein HP03

12	286	14.0	1308	22	AAF28689	Human protein HP03
13	257.8	12.6	617	21	AAA44349	Human secreted exp
14	193.2	9.5	572	21	AAA43744	Mouse secreted exp
15	127.6	6.2	936	22	AAF58252	Oligonucleotide D1
16	127.6	6.2	936	22	AAF58254	Oligonucleotide D1
17	127.6	6.2	936	22	AAF58257	Oligonucleotide D1
18	127.6	6.2	936	22	AAF58259	Oligonucleotide D2
19	127.6	6.2	936	22	AAF58262	Oligonucleotide D2
20	127.6	6.2	936	22	AAF58265	Oligonucleotide D2
21	122	6.0	936	22	AAF58252	Oligonucleotide D1
22	122	6.0	936	22	AAF58254	Oligonucleotide D1
23	122	6.0	936	22	AAF58257	Oligonucleotide D1
24	122	6.0	936	22	AAF58259	Oligonucleotide D2
25	122	6.0	936	22	AAF58262	Oligonucleotide D2
26	122	6.0	936	22	AAF58265	Oligonucleotide D2
27	100	4.9	1718	21	AAZ34958	Soybean acid triac
28	89.4	4.4	1483	21	AAZ34956	Rice acid triacylg
29	84.8	4.1	859	21	AAZ34953	Corn acid triacylg
30	72	3.5	1438	21	AAZ34959	Soybean acid triac
31	58.6	2.9	244	22	AAF58238	Oligonucleotide D1
32	58	2.8	244	22	AAF58238	Oligonucleotide D1
33	55.8	2.7	705	21	AAZ34952	Catalpa acid triac
34	54	2.6	549	21	AAZ34952	Cat flea hindgut a
35	51	2.5	509	21	AAZ34954	Corn acid triacylg
36	48.2	2.4	178	19	AAV35431	Cell death regulat
37	46.4	2.3	43676	21	AAZ60606	Nucleotide sequenc
38	42	2.1	8041	17	AAV10105	Adherence conferr
39	40.4	2.0	1140	20	AAZ04759	DNA encoding a coc
40	40.4	2.0	1857	22	AAZ59142	Human secreted pro
41	40.4	2.0	10628	22	AAZ57301	P. falciparum FC83
42	40.2	2.0	2021	19	AAV17087	Homo sapiens tubul
43	40	2.0	1937	21	AAZ47445	Macaque TANGO 232
44	40	2.0	2019	21	AAZ47445	Lung cancer associ
45	39.4	1.9	3189	20	AAZ20284	Borrelia burgdorfe

ALIGNMENTS

RESULT 1
AAF45131
ID AAF45131 standard; cDNA; 2044 BP.
XX
AC AAF45131.
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 cDNA.
XX

KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX

XX 24-MAY-2000; 2000WO-US14858.
XX
PR 14-JUN-1999; 99US-0333159.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX WPI; 2001-032313/04.
XX
DR P-PSDB; AAB66065.
XX

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
XX disease -

PS Claim 1: Fig 6; 359pp; English.

XX
CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.

XX Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;

Query Match 100.0%; Score 2044; DB 22; Length 2044;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	gtcgaccacgcgctcggggaattgcagcaggaataatggaagattgaaagattttaaaccac	60
DB	1	gtcgaccacgcgctcggggaattgcagcaggaataatggaagattttaaaccac	60
QY	61	aaattcttctttagaattagttgtgtacattggcaggaaaaataaataatgcagatgtt	120
DB	61	aaattcttctttagaattagttgtgtacattggcaggaaaaataaataatgcagatgtt	120
QY	121	ggaccatgttgaaacctgtcaagacagtgagttgtctcacacagaatggaaatgtgc	180
DB	121	ggaccatgttgaaacctgtcaagacagtgagttgtctcacacagaatggaaatgtgc	180
QY	181	ttctgattcttggcggtatattccagagaaatgtaattcagtacatatgccaaacta	240
DB	181	ttctgattcttggcggtatattccagagaaatgtaattcagtacatatgccaaacta	240
QY	241	aagctgtgaccacagaagcattcatgaaattagtgaaatcccaacatcaaggctatc	300
DB	241	aagctgtgaccacagaagcattcatgaaattagtgaaatcccaacatcaaggctatc	300
QY	301	cctgtgaggaatatgaatgcgaactgaagatgggtatattcctttctgttaacagattc	360
DB	301	cctgtgaggaatatgaatgcgaactgaagatgggtatattcctttctgttaacagattc	360
QY	361	ctcgaggcctagtgcaacctgaagacacaggttccaggcgtgtgttactgcagcatg	420
DB	361	ctcgaggcctagtgcaacctgaagacacaggttccaggcgtgtgttactgcagcatg	420
QY	421	gcctagtggagtgtagcaactggaattcccaacctgcccaacaatagcctggggtcca	480
DB	421	gcctagtggagtgtagcaactggaattcccaacctgcccaacaatagcctggggtcca	480
QY	481	ttctggcagatgctgtttttgacgtgtgtgaggggaacagcaggggaacgctgtctc	540
DB	481	ttctggcagatgctgtttttgacgtgtgtgaggggaacagcaggggaacgctgtctc	540
QY	541	gaaacacacagacactctccatagaccagaatgagttctgggctttcagttatgatgaa	600
DB	541	gaaacacacagacactctccatagaccagaatgagttctgggctttcagttatgatgaa	600
QY	601	tggctaggtttgacctctctgcagtgataaactttattttgcagaaaaacggccaggaaa	660
DB	601	tggctaggtttgacctctctgcagtgataaactttattttgcagaaaaacggccaggaaa	660
QY	661	agatctattatgtcgtctatcacaggccacacacatgggctttatgtcatcttccacca	720
DB	661	agatctattatgtcgtctatcacaggccacacacatgggctttatgtcatcttccacca	720

DB	661	agatctattatgtcgtctatcacaggccacacacatgggctttatgtcatcttccacca	720
QY	721	ttccagagctgctcagaaaaataaataatttttctttagcaccatagccacagtta	780
DB	721	ttccagagctgctcagaaaaataaataatttttctttagcaccatagccacagtta	780
QY	781	agcatgcaaaaaagccccgggaccacaaatttttgtgtgcagatatgatcaaggat	840
DB	781	agcatgcaaaaaagccccgggaccacaaatttttgtgtgcagatatgatcaaggat	840
QY	841	tgtttggcaaaaaagaatttctgtatcacaccagatttctcagaaactgttattacc	900
DB	841	tgtttggcaaaaaagaatttctgtatcacaccagatttctcagaaactgttattacc	900
QY	901	tttgtggcagtgattcttgcagattgttagtaataatcatgttacttctgggtgat	960
DB	901	tttgtggcagtgattcttgcagattgttagtaataatcatgttacttctgggtgat	960
QY	961	tcaacacacaaatatgaacatgagccgagcaagtgtatgtgtgccacactctgtctg	1020
DB	961	tcaacacacaaatatgaacatgagccgagcaagtgtatgtgtgccacactctgtctg	1020
QY	1021	gaacatctgtcaaaaatatctacactgagccagcagtgaaattcttctgtgaactccgg	1080
DB	1021	gaacatctgtgtcaaaaatatctacactgagccagcagtgaaattcttctgtgaactccgg	1080
QY	1081	catttgactgggggagtgagacacaaatctggaaaaatcgaaaaatcgaaactccctgtaa	1140
DB	1081	catttgactgggggagtgagacacaaatctggaaaaatcgaaaaatcgaaactccctgtaa	1140
QY	1141	ggtacagagtcagagatatgacggtccctacagcaatgtggacagaggtcagagactggc	1200
DB	1141	ggtacagagtcagagatatgacggtccctacagcaatgtggacagaggtcagagactggc	1200
QY	1201	tttcaaatccagaagcgtgaaatgtctctctctgaggtgacacacactcatctacata	1260
DB	1201	tttcaaatccagaagcgtgaaatgtctctctctgaggtgacacacactcatctacata	1260
QY	1261	agaaattctctgaatgggtccacgtggaattctctctgggttttggatgctccctcacgta	1320
DB	1261	agaaattctctgaatgggtccacgtggaattctctctgggttttggatgctccctcacgta	1320
QY	1321	gttacaatgaaatccatccatctgagtcagcagagagagacacacacttcccagggaggt	1380
DB	1321	gttacaatgaaatccatccatctgagtcagcagagagagacacacacttcccagggaggt	1380
QY	1381	gtgagggcgtattgtgaagcatctgacactgacatcttaggacaaactcctctgaggatg	1440
DB	1381	gtgagggcgtattgtgaagcatctgacactgacatcttaggacaaactcctctgaggatg	1440
QY	1441	gggctagggaccatgaagcagaaattacagagagcagagacactagatatttttcaga	1500
DB	1441	gggctagggaccatgaagcagaaattacagagagcagagacactagatatttttcaga	1500
QY	1501	ttcccgtgacttggcactaaatccgacacttacatttattttttctgtgtaataataa	1560
DB	1501	ttcccgtgacttggcactaaatccgacacttacatttattttttctgtgtaataataa	1560
QY	1561	gtacttattaggtaaatagaggttttgttgtctattatattctaccactcttgaagggt	1620
DB	1561	gtacttattaggtaaatagaggttttgttgtctattatattctaccactcttgaagggt	1620
QY	1621	aggttttaccctgatagccagaaaaatctcagacattctctatcatcattcaggtaaatctc	1680
DB	1621	aggttttaccctgatagccagaaaaatctcagacattctctatcatcattcaggtaaatctc	1680
QY	1681	tttaaaacacctctgttttttctataagccataatttttggagcactaaagtaaatggc	1740
DB	1681	tttaaaacacctctgttttttctataagccataatttttggagcactaaagtaaatggc	1740
QY	1741	aaatttgggacagatatgtgaggtctgtgagctgtggattattgttgcactttgacaaaataa	1800
DB	1741	aaatttgggacagatatgtgaggtctgtgagctgtggattattgttgcactttgacaaaataa	1800

QY 1801 gctagacattttcaacctgttgcacagagacataaacactacctcaggaagctgagctgc 1860
 Db 1801 gctagacattttcaacctgttgcacagagacataaacactacctcaggaagctgagctgc 1860
 QY 1861 tttaaggacaacaacaacaaatcagttgttacagtagatgagaaatctatttaagcatt 1920
 Db 1861 tttaaggacaacaacaacaaatcagttgttacagtagatgagaaatctatttaagcatt 1920
 QY 1921 ctcaagaataaggccaagttttatagttgcattctcagggaagaaaaattttataggaatttt 1980
 Db 1921 ctcaagaataaggccaagttttatagttgcattctcagggaagaaaaattttataggaatttt 1980
 QY 1981 atgagtttccaataaaatgcattctgcattacataaaaaaaataaaagggcgg 2040
 Db 1981 atgagtttccaataaaatgcattctgcattacataaaaaaaataaaagggcgg 2040
 QY 2041 ccgc 2044
 Db 2041 ccgc 2044

RESULT 2

AAF45132
 ID AAF45132 standard; CDNA; 1269 BP.

XX AAF45132;

DT 30-MAR-2001 (first entry)

XX Human TANGO 294 ORF.

DE Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.

XX Homo sapiens.

XX WO200077239-A2.

PD 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

DR P-PSDB; AAB66065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease.

XX Claim 1; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,

CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence: 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

Query Match 62.1%; Score 1269; DP:22; Length 1269;

Best Local Similarity 100.0%; Pred.No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 atgttggaaaaccttgcagacagtggtgtctcacacagagatgaaatgtggcttcg 185

Db 1 atgttggaaaaccttgcagacagtggtgtctcacacagagatgaaatgtggcttcg 60

QY 186 atctctgggtggcgatatgttccagagaaaatgtgaattcagtagacatgcacactaaagct 245

Db 61 atctctgggtggcgatatgttccagagaaaatgtgaattcagtagacatgcacactaaagct 120

QY 246 gtggaccacagaagcattcatgaattatgtgaaaatcccaacataaagcctatcctgt 305

Db 121 gtggaccacagaagcattcatgaattatgtgaaaatcccaacataaagcctatcctgt 180

QY 306 gaggaatatgaagtcgcgaactgaagatgggtatatctcttctgtaaacagagattcctga 365

Db 181 gaggaatatgaagtcgcgaactgaagatgggtatatctcttctgtaaacagagattcctga 240

QY 366 ggcctagtgcacacctaaagaacaggttccagggcctgtggttactgcagcatggccta 425

Db 241 ggcctagtgcacacctaaagaacaggttccagggcctgtggttactgcagcatggccta 300

QY 426 gtggaggtgtgtagcaactggaattcccaacctgcccacaaatagcctgggcttcattctg 485

Db 301 gtggaggtgtgtagcaactggaattcccaacctgcccacaaatagcctgggcttcattctg 360

QY 486 gcagatgctggtttgacgtgtggtggaacacagcaggggaaacgctggtctcga 545

Db 361 gcagatgctggtttgacgtgtggtggaacacagcaggggaaacgctggtctcga 420

QY 546 cacaagacactctccatagaccaagatgagttctgggttccagttatgatgagatggct 605

Db 421 cacaagacactctccatagaccaagatgagttctgggttccagttatgatgagatggct 480

QY 606 aggtttgacctctgcagtgataaacctttatttgcagaaaacggggccagaaaagatc 665

Db 481 aggtttgacctctgcagtgataaacctttatttgcagaaaacggggccagaaaagatc 540

QY 666 tattatgtcgctattccacagggccaccacatggcctttatgtcatttccaccatgcga 725

Db 541 tattatgtcgctattccacagggccaccacatggcctttatgtcatttccaccatgcga 600

QY 726 gagctggctcagaaaaatcaaaatgtatttggcttttagcaccatagccactgttaagcat 785

Db 601 gagctggctcagaaaaatcaaaatgtatttggcttttagcaccatagccactgttaagcat 660

QY 786 gcaaaaaagcccgagcaaaattttgtgtgcagatatgatgatcaagggatgttt 845

Db 661 gcaaaaaagcccgagcaaaattttgtgtgcagatatgatgatcaagggatgttt 720

QY 846 ggcacaaaagaattctctatcagaccagatttctcagacaaactgtttattacattgt 905

Db 721 ggcacaaaagaattctctatcagaccagatttctcagacaaactgtttattacattgt 780

QY 906 ggcaggtgattcttgatcagattttagtaataatcatgttacttctgggtggattcaac 965

Db 781 ggcaggtgattcttgatcagattttagtaataatcatgttacttctgggtggattcaac 840

QY 966 accacataatgaacatgacgagcaagtgatatgtgcacacactcttctgtgaaaca 1025

Db 841 accacataatgaacatgacgagcaagtgatatgtgcacacactcttctgtgaaaca 900

QY 1026 tctgtgcaaaatattctacactggcagcaggaattctgtgaaactccgggcat 1085

Db 901 tctgtgcaaaatattctacactggcagcaggaattctgtgaaactccgggcat 960

Db	1036	tgaccgccaatgagtcgaattgcagctgtggagtggtggcaatgactgtgttgctgacc	1095
QY	1210	cagaagacgtgaaatgctgctctctgagtgacaaacctcatctaccataagaatattc	1369
Db	1096	cccaagatgtgacctttgtctccaaacctttctaactcatctaccacaaaggagattc	1155
QY	1270	ctgaatgggctcacgtgattcatctctgggtttggatgctctccacgctatgtacaaatg	1329
Db	1156	ttccataataatcaccttgattttatctctgggcaatgaatgctctccagaagtttacaatg	1215
QY	1330	aaataatccatctgatgcagcaggaggaga	1359
Db	1216	aaattatttctatgatgtgcaaaaagataaaa	1245

RESULTS

RESULT 4
AAN60685

AAN000003
ID AAN60685 standard; DNA; 1367 BP.

XX
XX
XXXXXX

AC AAN60685;

XX

DT 08-JUN-1991 (first entry)

XX XX

DE Sequence encoding pregastric lipase.

XX

KW Precursor pol:

XX

OS Homo sapiens.

XX

EH	Key	-----
EH	Key	-----

FT . sig_peptide

E **B**

mat_peptide

XXIX

AX PN W08603778-A

FN XX
W08003776-A.

03-JUL-1986.

XX
27

PF 23-DEC-1985;

XX

PR 21-DEC-1984;

PR 23-DEC-1985;

PR 01-JAN-1986;

XX

PA (BREW-) BREWI

PA (TUBB/) TUBB

XX

PI Tubb RS;
.....

XX

DR WPI; 1986-182

DR P-PSDB; AAP60
yy

XX DT No: 00000000

PT New precursor used to trans-

used to trans
PT
yy

XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 XX
 PS
 XX
 XX Claim 2; Fig 1; 130pp; French.
 CC This is the nucleotide sequence encoding the dog pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 12 bases of this sequence (encoding amino acids
 CC 1-4 of the protein encoded by this sequence) or bases 1-162 of this
 CC sequence (encoding amino acids 1-54) can be deleted to form the
 CC derivatives designated delta-4 or delta-54 respectively. Plants, or
 CC their extracts, expressing the lipases or the truncated derivatives, can
 CC be used: (a) as pharmaceuticals or food to facilitate absorption of fat,
 CC either in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat absorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX
 SQ Sequence 1528 BP; 392 A; 387 C; 329 G; 420 T; 0 other;

Query Match 21.7%; Score 444.2; DB 17; Length 1528;
 Best Local Similarity 62.8%; Pred. No. 5.1e-104;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 acccagaagcattcgaatattagtagtaaatcatccaacatcaaggtatccctgtgagg 309
 DB 26 accctgaagtacataataagtcagatgatcatcactactgggagaccacccagctgagg 85
 QY 310 aatatgaagtcgaactgaagatgggtatatctcttctgttaacaggattccctcgagcc 369
 DB 86 aatatgaagtgtgacgaagacgggttatctcttgggagtcgacagaattcccttatggga 145
 QY 370 tagtgaacctgaagaacagaggttccagcctgtgtgttactgcagcagtgccctagttg 429
 DB 146 ggaataattcagagaataataggccggagacctgttgcatatttgcacacgggtttctcg 205
 QY 430 gagggtgtagaactgatttccaaactgcacaaatagcctgggtcttattcttggcag 489
 DB 206 catcagccaaactgatttccaaactgcacaaacagcctggccttcttctgcccgcg 265
 QY 490 atgctgttttgacgtgtgtagtgggaacagcaggggaaacgcctggtctcgaacaaca 549
 DB 266 acgcgggttacgcagctgt 325
 QY 550 agacactctccatagacacagatgagttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 609
 DB 326 tgtactactgcgcacactcgtcgaattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 385
 QY 610 ttgacctcttcgtagtgaataatttttttgcagaaaacggcggcaggaagatctatt 669
 DB 386 atgacctcttcgtagtgaataatttttttgcagaaaacggcggcaggaagatctatt 445
 QY 670 atgtcgtctattcacagggcaccacacatggggtttttgtgattttccaccatgcagagc 729
 DB 446 acgttggccattccagggccaccacacatggggtttttgtgattttccaccatgcagagc 505
 QY 730 tggctcagaaatacaaaatattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 789
 DB 506 tggcgaacggatcaaaacatttctgcatgtagctccgttgcacgggaagatcacagc 565
 QY 790 aagccccgggacaaaattttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 849
 DB 566 aaacctgttaacaaactcgtcgtcccttcttcttcttcttcttcttcttcttcttctt 625
 QY 850 aaaaagaatttctgtatcagaccagattttctcagacaaactgtttattacctttgtggcc 909

DB 626 acaaaatattaccacacacacactttttgatcaatttctgcaccaggtatgtccc 685
 QY 910 aggtgattttgatcagattttagtaataatcatgttacttcttgggttgattcaacaca 969
 DB 686 gcgagacggtgattctctctcctgcagcaacgcccgtttatcatcttgggattgacacta 745
 QY 970 acaatatgaacatgagccgagcaagtgatatgtctgtccacacacttctgtgggaacatctg 1029
 DB 746 tgaacttgaaactgagtcgcttggaatgtgtatctgtcacataatccagcaggaaatcgg 805
 QY 1030 tgcuaaattttacactgagcagcagcagtggaattctgtgtaactccggggtgact 1089
 DB 806 ttcagaaactgtctccactgtccaggtgttaagtctgggaagtctccaaagctttgact 865
 QY 1090 gggggagtgagaccaaaatctggaaaaatgcaatcagcacaactcctgttaaggtacagag 1149
 DB 866 gggggaagccagtcagaaatgagtcactatcatcagagcatgcccctactacaacc 925
 QY 1150 tcagagatatgacggtccctacagcaatgtggacagaggtgcaggactggctttcaaatc 1209
 DB 926 tgacagacatgcatgtgccaatcgcatgtggaaacggtggcaacgactgtggtggcagc 985
 QY 1210 cagagacgtgaaaatgctgtctctgtgagtgacaaacctctatcaccataagaatattc 1269
 DB 986 ctcccgatgtgaccttggcttctccaaagctccccaatctcattaccacaggaagattc 1045
 QY 1270 ctgcatgggctcagctggatttcatctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1329
 DB 1046 ctcttacaatacacttggactttatctgtggcattgagtcgccctcgaagcgtttacaatg 1105
 QY 1330 aaatcatccatctgatg 1346
 DB 1106 aaattgttccatgatg 1122

RESULT 10

AAQ68389
 ID AAQ68389 standard; DNA; 1531 BP.

XX AC AAQ68389;
 XX DT 20-FEB-1995 (first entry)
 XX DE Canine gastric lipase (Met-CGL).
 XX KW Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
 XX KW bioconversion; exocrine pancreatic insufficiency; ss.
 XX OS Canis familiaris.
 XX FH Key Location/Qualifiers
 XX FT CDS 1..1143
 XX FT /*tag- a
 XX PN W09413816-A.
 XX PD 23-JUN-1994.
 XX PF 16-DEC-1993; 93WO-FR01260.
 XX PR 16-DEC-1992; 92FR-0015201.
 XX PA (LJOU) INST RECH JOUVEINAL.
 XX PI Benicourt C, Blanchard C, Junien J;
 XX DR WPI; 1994-217890/26.
 XX DR P-PSDB; AAR56871.
 XX PT Recombinant canine gastric lipase and nucleic acid encoding it -
 XX PT are used for improving absorption of ingested fat, treating
 XX PT mucoviscidosis etc. and in enzymatic bio-conversions

XX PS Claim 2; Fig 8; 52pp; French.

XX CC The sequence given below is the sequence of figure 8, altered

CC according to the amendments described on page 2 of the appended

CC letter.

CC CGL is used to improve absorption of ingested fat, in healthy and

CC sick patients (e.g. having altered levels of gastric lipase); to

CC treat conditions associated with insufficiency (or lack) of lipases,

CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.

CC where immobilised, for bioconversions, e.g. hydrolysis or

CC transesterification (other mammalian gastric lipases, or derivs.,

CC can be used in this application).

XX CC Sequence 1531 BP; 395 A; 386 C; 329 G; 421 T; 0 other;

XX SQ

Query Match 21.7%; Score 444.2; DB 15; Length 1531;

Best Local Similarity 62.8%; Pred. No. 5.1e-104;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 250 acccagaagcattcatgaatattagtgaaatcatcccaacatcaaggctatccctgtgagg 309

Db 29 accctgaagtgaccatgaatataagtcagatgacacactactggggataccacagtgagg 88

Qy 310 aatatgaagtcgaactgaagatgggtatatcccttttctgttaacaggatccctcgagcc 369

Db 89 aatatgaagttgtgaccgaagacggttatctcttggatcgacagaattccttatggga 148

Qy 370 tagtgcacacctaaagaacaggtcccgacctgtgtgttactgcagcagcatggcctagt 429

Db 149 ggaataattcagaaataataggccgagacactgtgtgatattttgcaacaggtttgctg 208

Qy 430 gaggtgtagcaactgatttcccaacctgcccacataagctagctgggttcttattctgag 489

Db 209 catcagccacaactgattctccacctgcccacacacagctggccttctatctctgccc 268

Qy 490 atgtcgttttgacgtgtgagtggaacagcagcaggggaacgcctgtgtctcgaacaca 549

Db 269 acgcgcgggtacgcgtgtgctgggggaacagcaggggaacacacctggccaggaggatc 328

Qy 550 agacactctccatagaccagatgattctgggttcttcagttatgatgagatggctaggt 609

Db 329 tgtactactgcgcgactccgtcgaattctgggttcttcagctttgacgagatgctaat 388

Qy 610 ttgaacctctgcagtgataaactttattttgcagaacacggccaggaagaaatctatt 669

Db 389 atgaccttccgcacacattgacttcatcttgaagaaacggggacaggaacagctacact 448

Qy 670 atgtcggctattcagggacacacacacatggtttatgttatttccaccatgcagagc 729

Db 449 acgttggccattccagggacacacacattggttctcgccttttccaccaatcccaagc 508

Qy 730 tggctcagaaatacaaaatgattttgttcttcagccacacatagccactgttaagcatcaa 789

Db 509 tggcgaacggatcaaaacctctatgcattagctccgttgcacacgtgaagtaacccg 568

Qy 790 aaagcccccggaccacaaattttttgtgtgcagatagatgatacgaaggatgttggtgca 849

Db 569 aaacctgttaaaacaaactcgtctccttcttcttcttcccttcaagcttatattggaa 628

Qy 850 aaaaagaattctgtatcagacacagatttctcagacacactgttatttatttctgtggcc 909

Db 629 acaaaatttctacccacacacacttcttctgtatcatatttctgcgacacgaggtatgctccc 688

Qy 910 aggtgattcttgatcagattgttagtaatacatcatgttacttctggtggattcaacacca 969

Db 689 gcgagacggttgatctctctcgtcagacacgcctcgtttatatttggatttgacacta 748

Qy 970 acaatatgaacatgagccgagcagtgatatgtctgccacactcttctgtgggaacatctg 1029

Db 749 tgaacttgaacatgagtcgtgtgtgtatctgttcacataatccacagcaggaacatcgg 808

Qy 1030 tgcaaaatattctacactggagccagcagcagtgatattctgtgaactccggcgtttgaact 1089

Db 809 tcagaaacgtgtccactgtgtccagcagcgttgaagtctggaagttccaagcttttgact 868

Qy 1090 gggggagtgagaccacaaatctgaaaaatgcaaatcagcgaactcctgttaagggtacagag 1149

Db 869 ggggaagccagtcagaaacatgatgcactatcatcagagcagcctcctcctactacaacc 928

Qy 1150 tcagagatatgacggtccctacacagcaatgtggacagagcagtgactggctttcaaatc 1209

Db 929 tgacagacatgcagtggccaatcgagctgtggaacggtgcaacgacttctggtccgacg 988

Qy 1210 cagaagacgtgaaatctgctctctctgtgagtgacccaacctcatcaccataagaatttc 1269

Db 989 ctacagatgttgacctttgtcttccaagctccccaatctcatttaccacagggagattc 1048

Qy 1270 ctgaaatgggctcagctgggatttctctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1329

Db 1049 ctcttacaatcacttggactttatcttggccatggatgcccctcaagcgttttacaatg 1108

Qy 1330 aaatcatcctctgtatg 1346

Db 1109 aaattgttccatgatg 1125

RESULT 11

AAF28679

ID AAF28679 standard; cDNA; 699 BP.

XX AC AAF28679;

XX XX

DT 05-APR-2001 (first entry)

XX XX

DE Human protein HP03372 coding sequence #1.

XX XX

KW Human; hydrophobic domain; Immune deficiency; autoimmune disorder;

KW allergy; tissue growth; regeneration; wound healing; burn; tumour;

KW periodontal disease; thrombolytic condition; haemostatic condition;

KW infection; ss.

XX XX

OS Homo sapiens.

XX XX

PN WO200102563-A2.

XX XX

PD 11-JAN-2001.

XX XX

PF 16-JUN-2000; 2000WO-JP03943.

XX XX

PR 02-JUL-1999; 99JP-0188835.

XX XX

PA (SAGA) SAGAMI CHEM RES CENT.

PA (PROT-) PROTEGENE INC.

XX XX

PI Kato S, Kimura T;

XX XX

WP: 2001-071581/08.

DR P-PSDB; AAB61608.

XX XX

XX New human proteins with hydrophobic domains, useful for the treatment

PT of immune disorders, tumors, allergic conditions, thrombosis and

PT microbial infection

XX XX

PS Cialm 3; Page 113; 153pp; English.

XX XX

CC The present invention relates to human proteins (AAB61608-AAB61617) and

CC their coding sequences (AAF28679-AAF28698). The proteins of the present

CC invention have hydrophobic domains and can be used for the treatment of

CC various immune deficiencies and disorders, such as severe combined

CC immunodeficiency (SCID), multiple sclerosis, rheumatoid arthritis,

CC autoimmune pulmonary inflammation, graft-versus-host disease and

CC Guillain-Barre syndrome. The proteins may also be useful in the treatment

CC of allergic reactions and conditions, such as asthma and in regulation of

CC haematopoiesis or lymphoid cell deficiencies. The proteins may also have

CC utility in compositions used for bone, cartilage, tendon and/or nerve
CC tissue growth or regeneration as well as wound healing and in the
CC treatment of burns. The proteins may be used in the treatment of
CC periodontal disease and in other tooth repair processes. Other uses
CC include treatment of thrombolytic and haemostatic conditions, treatment
CC or prevention of tumours and inhibiting infection by bacteria, viruses,
CC fungi and other parasites.

XX Sequence 699 BP; 216 A; 144 C; 148 G; 191 T; 0 other;

Query Match 14.0%; Score 286; DB 22; Length 699;

Best Local Similarity 67.4%; Pred. No. 1.2e-63;

Matches 403; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 211 gaaatgtgaattcagtcacatgccaaactaaagctgtggaccacagagcattcatgaata 270
DB 44 gatctatgtgtgttatgacaaaggaagaaacatgcaaacctggaagcctaataatgaata 103
QY 271 ttagtgaatccatccacatacaaggctatccctgtgaggaaatgaagtcgcaactggaag 330
DB 104 ttaccagattattcttactggtttatcttcttgaagagtgatgatttacaacaaaag 163
QY 331 atgggtatattcttctgttaacagagattctctgagccttagtgcacctaagaagacag 390
DB 164 atgggtatattcttctgttaacagagattctctgagccttagtgcacctaagaagacag 223
QY 391 gtccagcctgtgttactcagcagctgcttagtgcagagtgctgagagtgctgacactgatt 450
DB 224 ctcaaacgctgtgttattgcagcagctgcttattgcatctgcatctgcatctgatt 283
QY 451 ccaacctgcccaacatagctgctgcttcttctgagagatgctgtgttgacgtgtgga 510
DB 284 gcaacctgcccaacacagttgttcttcttctgagagatgctgtgttgacgtgtgga 343
QY 511 tggggacacagcaggggaacagcctgtgtctgaaacacacagacacttccatgacacaa 570
DB 344 tggggacacagcaggggaacacagcctgtgtctgaaacacacacttccatgacacaa 403
QY 571 atgagttctggtcttcagttatgagatggtgtgaggtgtgaccttccctgagtgataa 630
DB 404 cagaatactggccttcagttgtgagatggtgtgaggtgtgaccttccctgagtgataa 463
QY 631 acttatttgcagaaacagggccaggaagaaatctattatgtcgtgtatttcacaggga 690
DB 464 atttattcagaaacagggccaggaagaaatctattatgtcgtgtatttcacaggga 523
QY 691 ccacatggccttattgtatttccacatgacagagctggtgtgagaaatcaaaatgt 750
DB 524 ccacatggccttattgtatttccacatgacagagctggtgtgagaaatcaaaatgt 583
QY 751 atttggcttgcacccatagcactgttgaagcagcgaagaaagcccggaacaaatt 808
DB 584 ttttgcactggctcagttgtcacagtttaataacaccccaaaagtccttatgaaaaaact 641

RESULT 12

AAF28689

ID AAF28689 standard; cDNA; 1308 BP.

XX AAF28689;

AC AAF28689;

XX 05-APR-2001 (first entry)

DT Human protein HP03372 coding sequence #2.

DE Human; hydrophobic domain; immune deficiency; autoimmune disorder;

XX allergy; tissue growth; regeneration; wound healing; burn; tumour;

KW periodontal disease; thrombolytic condition; haemostatic condition;

KW infection; ss.

XX Homo sapiens.

OS

XX

PN WO200102563-A2.

XX 11-JAN-2001.

PD 16-JUN-2000; 2000WO-JP03943.

XX 02-JUL-1999; 99JP-0188835.

XX (SAGA) SAGAMI CHEM RES CENT.

XX (PROT-) PROTEGENE INC.

PA Kato S, Kimura T;

PI WPI; 2001-071581/08.

XX P-PSDB; AAB61608.

XX New human proteins with hydrophobic domains, useful for the treatment

PT of immune disorders, tumors, allergic conditions, thrombosis and

PT microbial infection -

XX Claim 4; Pages 122-124; 153pp; English.

XX The present invention relates to human proteins (AAB61608-AAB61617) and

CC their coding sequences (AAF28679-AAF28698). The proteins of the present

CC invention have hydrophobic domains and can be used for the treatment of

CC various immune deficiencies and disorders, such as severe combined

CC immunodeficiency (SCID), multiple sclerosis, rheumatoid arthritis,

CC autoimmune pulmonary inflammation, graft-versus-host disease and

CC Guillain-Barre syndrome. The proteins may also be useful in the treatment

CC of allergic reactions and conditions, such as asthma and in regulation of

CC haematopoiesis or lymphoid cell deficiencies. The proteins may also have

CC utility in compositions used for bone, cartilage, tendon and/or nerve

CC tissue growth or regeneration as well as wound healing and in the

CC treatment of burns. The proteins may be used in the treatment of

CC periodontal disease and in other tooth repair processes. Other uses

CC include treatment of thrombolytic and haemostatic conditions, treatment

CC or prevention of tumours and inhibiting infection by bacteria, viruses,

CC fungi and other parasites.

XX Sequence 1308 BP; 402 A; 246 C; 264 G; 396 T; 0 other;

Query Match

Best Local Similarity 67.4%; Pred. No. 1.5e-63;

Matches 403; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 211 gaaatgtgaattcagtcacatgccaaactaaagctgtggaccacagagcattcatgaata 270
DB 119 gatctatgtgtgttatgacaaaggaagaaacatgcaaacctggaagcctaataatgaata 178
QY 271 ttagtgaatccatccacatacaaggctatccctgtgaggaaatgaagtcgcaactggaag 330
DB 179 ttaccagattattcttactggtgttactcttcttgaagagtgatgatttacaacaaaag 238
QY 331 atgggtatattcttctgtttaaagagattctctgagccttagtgcacctaagaagacag 390
DB 239 atgggtatattcttctgtttaaagagattctctgagccttagtgcacctaagaagacag 298
QY 391 gtccagcctgtgtgttactcagcagctggtgtgaggtgtgaggtgtgaggtgtgaggtgtt 450
DB 299 ctccaaagcctgtgtgttattgcagcagctgttattgcatctgcatctgcatctgaggttt 358
QY 451 ccaacctgcccaacataagcctgtgtgttcttcttcttcttcttcttcttcttcttcttctt 510
DB 359 gcaacctgcccaacataagcctgtgtgttcttcttcttcttcttcttcttcttcttcttctt 418
QY 511 tggggacacagcaggggaacacgctgtgtctgaaacacacagacacttccatgacacaaag 570
DB 419 tggggacacagcaggggaacacgctgtgtctgaaacacacactttaaattgtcaccggaatcac 478
QY 571 atgagttctggtgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 630
DB 479 cagaatactggccttcagttgtgagatggtgtgagatggtgtgagatggtgtgagatggtgt 538

QY 631 actttattttcagaaaaagggccaggaagaaagatctattatgtcggctatttcacagggca 690
 Db 539 attttatcatagaaaaagcaggaagcactctactactcgtggccactcacaaggca 598
 QY 691 ccacatgggtttatttcatttccaccatgcccagagctgctcagaaaaatcaaatgt 750
 Db 599 ccacatagcttttatagcattttctcaaacaccagactggctcaaaagattagatat 658
 QY 751 attttgcttagcaccatagcactgttaagcagcagaaagcccgggaccacaaat 808
 Db 659 tttttgcagctcctcaggtgtcagagttacagtttaaacaccacaaagccttatgaaaaaact 716

RESULT 13

AAAA4349
 ID AAA44349 standard; CDNA: 617 BP.

AC AAA44349;

DT 21-AUG-2000 (first entry)

DE Human secreted expressed sequence tag SEQ ID NO:924.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytosolic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; cytostatic; antiparkinsonian;
 KW antiulcer; osteopathic; neuroprotective; vulnary; antiparkinsonian;
 KW cerebrotective; anticonvulsant; antidepressant; gene therapy;
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
 KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
 KW central nervous system disorder; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; coagulation disorder;
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
 KW tumour; infection; depression; psoriasis; ss.

OS Homo sapiens.

XX WO200021991-A1.

XX 20-APR-2000.

XX 15-OCT-1999; 99WO-US24206.

XX 15-OCT-1998; 98US-010436.

XX (GEMV) GENETICS INST INC.

XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 XX Merberg D, Treacy M, Bowman MR;

XX WPI; 2000-317938/27.

PT Isolated polynucleotides, and encoded proteins, comprising secreted
 PT expressed sequence tags (SESTs), useful for treating various disorders
 PT such as autoimmune, infectious, and central nervous system disorders -

PS Claim 1; Page 441; 803pp; English.

XX AAA43426 to AAA5925 represent specifically claimed secreted expressed
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat
 CC tissue sources. The SESTs can have a range of activities depending on
 CC the tissues they were isolated from. The activities include:
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;
 CC antiasthmatic; vulnary; antiulcer; osteopathic; neuroprotective;
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebrotective;
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene
 CC therapy and in vaccines. The SESTs are useful as probes for the

CC identification and isolation of full-length cDNAs and genomic DNA
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs
 CC are useful in assays for determining biological activity and raising
 CC antibodies. They may be useful for treatment of autoimmune disorders
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, ulcers,
 CC osteoporosis, osteoarthritis, central nervous system disorders
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's
 CC disease), tumours, bacterial, fungal or viral infections, depression and
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given
 CC in the exemplification of the present invention.

XX Sequence 617 BP; 181 A; 125 C; 139 G; 172 T; 0 other;

Query Match 12.6%; Score 257.8; DB 21; Length 617;
 Best Local Similarity 70.7%; Pred. No. 1.8e-56;
 Matches 343; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 244 ctgaggaccaggaagcattcatgaattatagtgaaatcatccaacatcaaggatccct 303
 Db 133 ctggaagccctgaagtgaactgaacattagtcagatgattactattgggataccca 192
 QY 304 gtgaggatgaagtgcgaactgaagatgggtatatactcttctgttaacggattccct 363
 Db 193 atgagaatgaagtgtgactgaagatggtattattcttgagtcataagaattccct 252
 QY 364 gaggccttagtgcaacctaaagaagcaggttccaggcctgtgtgttactcagcagcc 423
 Db 253 atgggaagaaaaattcggggaatacaggccagagactgtgtgttttgcagcatggt 312
 QY 424 tagttgaggtgtcagcaactggatttccaaacctgcccaacatagcctggcttcattc 483
 Db 313 tgcctgatcagccacaaactggatttccaaacctgccgaacacagccttgcttcattc 372
 QY 484 tggcagatgctggttttgacgtgtggatggggaacagcaggggaaacccctggtctgaa 543
 Db 373 tggcagatgctggtttatgattgtgtgctggcagacagagagaaacacctgggcaga 432
 QY 544 aacacaagacactctccatagaccagatgattcttggccttcagttatgagatgg 603
 Db 433 gaaacttgactattaccagattcagttgaattcttggccttcagcttgatgaatgg 492
 QY 604 ctggtttgacctctcctgagtgataaaactttatttcagaaaaacggcaggaaga 663
 Db 493 ctgaatagaccttccagccacaatcgacttcattgttaagaaaaactggacagagcgc 552
 QY 664 tctattatgtcgctattcacagggccaccacatgggctttattgcatctttccaccatgc 723
 Db 553 tacactatgtggcattccagggccaccacattggtttattgctctttccaccatc 612
 QY 724 cagag 728
 Db 613 tcgag 617

RESULT 14

AAA43744

ID AAA43744 standard; CDNA: 572 BP.

AC AAA43744;

DT 21-AUG-2000 (first entry)

DE Mouse secreted expressed sequence tag SEQ ID NO:319.

XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;
 KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
 KW antiviral; antidiabetic; antiasthmatic; vulnary; antiparkinsonian;
 KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;

```

XX  The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX  Sequence 936.BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ

```

Search completed: August 24, 2001, 22:41:42
Job time: 7403 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 20:11:58 ; Search time 4395.77 Seconds
(without alignments)
7192.387 Million cell updates/sec

Title: US-09-333-159-45
Perfect score: 2044
Sequence: 1 gtcgaccacgcgtccgggg.....aaaaaaaaaaggcgccgc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_ba3:
4: gb_in1:
5: gb_in2:
6: gb_in3:
7: gb_om:
8: gb_ov:
9: gb_pat1:
10: gb_pat2:
11: gb_ph:
12: gb_pl1:
13: gb_pl2:
14: gb_pl3:
15: gb_pl4:
16: em_ba1:
17: em_ba2:
18: em_fun:
19: em_htgo_hum:
20: em_htgo_inv:
21: em_htgo_rod:
22: em_htg_hum1:
23: em_htg_hum2:
24: em_htg_hum3:
25: em_htg_hum4:
26: em_htg_hum5:
27: em_htg_hum6:
28: em_htg_hum7:
29: em_htg_hum8:
30: em_htg_inv1:
31: em_htg_inv2:
32: em_htg_other:
33: em_htg_rod:
34: em_hum1:
35: em_hum2:
36: em_hum3:
37: em_hum4:
38: em_hum5:
39: em_hum6:
40: em_hum7:
41: em_in:
42: em_om:
43: em_or:

44: em_ov:
45: em_pat:
46: em_ph:
47: em_pl:
48: em_ro:
49: em_sts:
50: em_sy:
51: em_un:
52: em_v1:
53: gb_sts1:
54: gb_sts2:
55: gb_sts3:
56: gb_sy:
57: gb_un:
58: gb_v11:
59: gb_v12:
60: gb_htg1:
61: gb_htg2:
62: gb_htg3:
63: gb_htg4:
64: gb_htg5:
65: gb_htg6:
66: gb_htg7:
67: gb_htg8:
68: gb_htg9:
69: gb_htg10:
70: gb_htg11:
71: gb_htg12:
72: gb_htg13:
73: gb_htg14:
74: gb_htg15:
75: gb_htg16:
76: gb_htg17:
77: gb_htg18:
78: gb_htg19:
79: gb_htg20:
80: gb_htg21:
81: gb_htg22:
82: gb_htg23:
83: gb_htg24:
84: gb_htg25:
85: gb_pr1:
86: gb_pr2:
87: gb_pr3:
88: gb_pr4:
89: gb_pr5:
90: gb_pr6:
91: gb_pr7:
92: gb_pr8:
93: gb_pr9:
94: gb_rol:
95: gb_rol2:
96: gb_in4:
97: gb_pr10:
98: em_ba3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	892	43.6	181170	79	AL353113
c 2	890.4	43.6	158195	63	AC015461 Homo sapi
c 3	630.2	30.8	76653	63	AC015506 Homo sapi
4	527.8	25.8	2481	93	HSLYACLY 231690 H.sapiens (
5	527.8	25.8	2481	97	HSU08464 Human lysos
6	527.8	25.8	2626	93	HSLAL U08464 Human lysos
7	526.2	25.7	2493	97	HUMLIPCHL X76488 H.sapiens m
8	489.2	23.9	1378	9	A26689 M74775 Human lysos
					A26689 Precursor o

```

c 9 489.2 23.9 1378 9 A26690
10 470 23.0 1365 93 HSGIR
11 470 23.0 1367 9 A01046
12 470 23.0 1367 9 A01046
13 467.6 22.9 1336 9 A12714
14 467.6 22.9 1336 9 A01157
15 461 22.6 2358 94 MMLVACLY
16 459.8 22.5 3144 95 S81497
17 458 22.4 1417 7 BOVME
18 444.2 21.7 1137 9 AR039022
19 444.2 21.7 1137 9 AR02633
20 444.2 21.7 1140 9 AR039023
21 444.2 21.7 1140 9 AR02634
22 444.2 21.7 1146 9 AR039024
23 444.2 21.7 1146 9 AR02635
24 444.2 21.7 1198 9 A57760
25 444.2 21.7 1528 9 A57756
26 444.2 21.7 1528 9 AR039021
27 444.2 21.7 1528 9 AR02632
28 444.2 21.7 1651 7 CFLIPASE
29 441 21.6 1137 9 A39301
30 441 21.6 1140 9 A39303
31 441 21.6 1146 9 A39305
32 441 21.6 1528 9 A39300
33 388.8 19.0 1048 9 A57758
34 286 14.0 699 10 AX068249
35 286 14.0 1308 10 AX068259
36 255 12.5 148720 65 AC019353
37 255 12.5 161826 90 AL358532
38 255 12.5 172834 62 AC011277
39 151.2 7.4 76653 63 AC015506
40 140.6 6.9 401 93 HSLIPCHL04
41 140.6 6.9 1851 93 HSLIP04
42 140.6 6.9 180366 81 AL513533
43 127.6 6.2 40114 6 CEL2K6
44 122.4 6.0 137911 61 AC009769
45 122.4 6.0 204102 63 AC013826

```

ALIGNMENTS

```

RESULT 1
AL353113 181170 bp DNA HTG 08-APR-2001
LOCUS Homo sapiens chromosome 10 clone RP11-30415, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 unordered pieces.
ACCESSION AL353113.4 GI:13567970
VERSION HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 181170)
Grafham, D.
Direct Submission
Submitted (07-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
On Apr 9, 2001 this sequence version replaced gi:9796988.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA30415
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 180405 bases at least Q40

```

```

Consensus quality: 180692 bases at least Q30
Consensus quality: 180914 bases at least Q20
Insert size: 181070; sum-of-contigs
Insert size: 150393; agarose-fp
Quality coverage: 7.50x in Q20 bases; sum-of-contigs Quality
coverage: 9.09x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 174783: contig of 174783 bp in length
* 174784 174883: gap of 100 bp
* 174884 181170: contig of 6287 bp in length.

```

FEATURES

source

```

1..181170
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-30415"
/clone_lib="RPC1-11.2"
1..174783
/note="assembly_fragment:02734
clone_end:SP6
vector_side:left"
174884..181170
/note="assembly_fragment:02008"

```

misc_feature

```

1..174783
/note="assembly_fragment:02734
clone_end:SP6
vector_side:left"
174884..181170
/note="assembly_fragment:02008"

```

misc_feature

```

174884..181170
/note="assembly_fragment:02008"

```

```

BASE COUNT 57065 a 34281 c 33093 g 56630 t 101 others
ORIGIN

```

```

Query Match 43.6%; Score 892; DB 79; Length 181170;
Best Local Similarity 100.0%; Pred. No. 2.2e-213;
Matches 892; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1126 agccaaactctgttaaggtacagagatgacagatgacggtccctacagcaatgagagacag 1185
|||||
DB 134716 AGCCAACTCTGTAAAGGTACAGAGATGACGCGTCCCTACAGCAATGAGACAG 134775
|||||
QY 1186 gaggtcaggactggcttcaaatccagagacgctgaaatgctgctctgagtgacca 1245
|||||
DB 134776 GAGGTCAGGACTGGCTTCAATCCAGAACGCTGAAATGCTGCTCTCTGAGGTGACCA 134835
|||||
QY 1246 acctcatcaccataagaataattccctgaatgggctacgctgagatttcattctgaggtttg 1305
|||||
DB 134836 ACCTCATCTACCAATAAGAAATATTCCTGAATGGGCTCACGTTGATTCATCTGGGGTTGG 134895
|||||
QY 1306 atgctctcaccgtatgataatgaaatccatccatctgagcagcagagagagacacacc 1365
|||||
DB 134896 ATGCTCTCTACCGTATGTACAATGAATCATTCATCTGATGACGAGAGAGACCAACC 134955
|||||
QY 1366 ttccccaggagacggttgagggccgtatttgaaagcatctgacactgacgtattcttaggaca 1425
|||||
DB 134956 TTTCCTCCAGGAGCGGTGTAGGGCGGTATTGTGAAGCATCTGCACACTGACGATCTTAGGACA 135015
|||||
QY 1426 acctctgaggggtggggttaggaccatgaagcagaataccagagagagagacacacag 1485
|||||
DB 135016 ACCTCTCTAGGGATGGGGCTAGGACCCCATGAAGCGAAATTCGGAGAGAGAGACACATAG 135075
|||||
QY 1486 tacaatttttcagattccctgcaactggcactaaatccgacactacattacattttt 1545
|||||
DB 135076 TATCATATTTTCAGATTCCTGCGACTGGCACTAAATCCACACTTACATTATTTT 135135
|||||
QY 1546 ttctgtaaattaaagtactatttaggttaaatagaggttttgatgctattatattct 1605
|||||
DB 135136 TTCTCTGAATTAAGTACTATTATAGGTAATAGAGGTTTGTGTATGCTATTATATATTTCT 135195
|||||
QY 1606 accatcttgaggggttaggttttaccctgataccagagagaaatattatagacattcttatc 1665
|||||

```

```

Db 135196 ACCATCTTTGAAGGGTAGCTTTTACCTGATAGCCAGAAAATATCTAGACATCTCTATATC 135255
QY 1666 attcaggtaaatctctttaaacacactattgtttttttataaagccatatttttggagca 1725
|||||
Db 135256 ATTCAAGTAAATCTCTTTAAACACACCTATTCTTTTCTATAGCCATATTTTGGAGCA 135315
|||||
QY 1726 ctcaagataaatggcaaatgggacagatatagggtctgagctctgagctctggtattgttg 1785
|||||
Db 135316 CTAAAGTAAATGGCAATTTGGACAGATATTGAGTCTGGAGTCTGGATTAATGTTG 135375
|||||
QY 1786 actttgacaaaataagctagacattttcaccttctgttcacagagacataacactacctc 1845
|||||
Db 135376 ACTTTGACAAAATAAGCTAGACATTTTTCACCTTGTGTCACAGACATACACTACCTC 135435
|||||
QY 1846 aggaagctgagctgttttaagagacaaacaaacaaacaaacaaacaaacaaacaaacaaac 1905
|||||
Db 135436 AGGAAGCTGAGCTGCTTTTAAAGACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 135495
|||||
QY 1906 tctatttaagcattctcagataagccaaagttttatagttgcatctcaggggaagaaaa 1965
|||||
Db 135496 TCTATTGTAAGCATCTCTCAGATAGGCCAAGCTTTTATGTTGCACTCAGGGAAGAAAA 135555
|||||
QY 1966 tttatagatgtttatgagttctccaaataaatgcattctgcattacataaa 2017
|||||
Db 135556 TTTTATAGATGTTTATGATGTTCTCCATAAATGATGCTGCATTACATATAA 135607
|||||

RESULT 2
AC015461/c
LOCUS
DEFINITION Homo sapiens clone RP11-115N17, WORKING DRAFT SEQUENCE, 10
ACCESSION AC015461
VERSION AC015461.3 GI:9502445
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-115N17
Unpublished
2 (bases 1 to 158195)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgaiter,B.,
Brown,A., Castle,A., Collinge,M., Collins,S., Collamore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferrelira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczký,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
resfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 27, 2000 this sequence version replaced gi:9112670.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L1907
Center clone name: 115_N_17

```

```

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152466 bases at least Q40
Consensus quality: 154897 bases at least Q30
Consensus quality: 156006 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 157295; sum-of-contigs
Quality coverage: 5.8 in Q20 bases; agarose-fp
Quality coverage: 5.8 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 1002: contig of 1002 bp in length
* 1003 1102: gap of 100 bp
* 1103 2388: contig of 1286 bp in length
* 2389 2488: gap of 100 bp
* 2489 3525: contig of 1037 bp in length
* 3526 3625: gap of 100 bp
* 3626 16527: contig of 12902 bp in length
* 16528 16627: gap of 100 bp
* 16628 28636: contig of 12009 bp in length
* 28637 28736: gap of 100 bp
* 28737 46598: contig of 17862 bp in length
* 46599 46698: gap of 100 bp
* 46699 71912: contig of 25214 bp in length
* 71913 72012: gap of 100 bp
* 72013 96612: contig of 24600 bp in length
* 96613 96712: gap of 100 bp
* 96713 128331: contig of 31619 bp in length
* 128332 128431: gap of 100 bp
* 128432 158195: contig of 29764 bp in length.

```

FEATURES

source

```

1. 158195
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-115N17"
/clone_lib="RPC1-11 Human Male BAC"

```

1. 1002

misc_feature

```

1103 2388
/note="assembly_fragment"

```

misc_feature

```

2489 3525
/note="assembly_fragment"

```

misc_feature

```

3626 16527
/note="assembly_fragment"

```

misc_feature

```

16628 28636
/note="assembly_fragment"

```

misc_feature

```

28737 46598
/note="assembly_fragment"

```

misc_feature

```

46699 71912
/note="assembly_fragment"

```

misc_feature

```

72013 96612
/note="assembly_fragment"

```

misc_feature

```

96713 128331
/note="assembly_fragment"

```

misc_feature

```

128432 158195
/note="assembly_fragment"

```

misc_feature

```

clone_end:SP6
vector_side:right

```

misc_feature

```

28737 46598
/note="assembly_fragment"

```

misc_feature

```

46699 71912
/note="assembly_fragment"

```

misc_feature

```

72013 96612
/note="assembly_fragment"

```

misc_feature

```

96713 128331
/note="assembly_fragment"

```

misc_feature

```

128432 158195
/note="assembly_fragment"

```

misc_feature

```

clone_end:SP6
vector_side:left

```

```

BASE COUNT 48628 a 30917 c 30188 g 47557 t 905 others
ORIGIN

```


*	10620	11417:	contig of 798 bp	in length
*	11418	11517:	gap of 100 bp	
*	11518	12296:	contig of 779 bp	in length
*	12297	12396:	gap of 100 bp	
*	12397	13180:	contig of 784 bp	in length
*	13181	13280:	gap of 100 bp	
*	13281	14065:	contig of 785 bp	in length
*	14066	14165:	gap of 100 bp	
*	14166	14949:	contig of 784 bp	in length
*	14950	15049:	gap of 100 bp	
*	15050	15922:	contig of 773 bp	in length
*	15823	15922:	gap of 100 bp	
*	15923	16718:	contig of 796 bp	in length
*	16719	16818:	gap of 100 bp	
*	16819	17621:	contig of 803 bp	in length
*	17622	17721:	gap of 100 bp	
*	17722	18517:	contig of 796 bp	in length
*	18518	18617:	gap of 100 bp	
*	18618	19390:	contig of 773 bp	in length
*	19391	19490:	gap of 100 bp	
*	19491	20253:	contig of 763 bp	in length
*	20254	20353:	gap of 100 bp	
*	20354	21134:	contig of 781 bp	in length
*	21135	21234:	gap of 100 bp	
*	21235	22028:	contig of 794 bp	in length
*	22029	22128:	gap of 100 bp	
*	22129	22905:	contig of 777 bp	in length
*	22906	23005:	gap of 100 bp	
*	23006	23787:	contig of 782 bp	in length
*	23788	23887:	gap of 100 bp	
*	23888	24648:	contig of 761 bp	in length
*	24649	24748:	gap of 100 bp	
*	24749	25529:	contig of 781 bp	in length
*	25530	25629:	gap of 100 bp	
*	25630	26414:	contig of 785 bp	in length
*	26415	26514:	gap of 100 bp	
*	26515	27308:	contig of 794 bp	in length
*	27309	27408:	gap of 100 bp	
*	27409	28169:	contig of 761 bp	in length
*	28170	28269:	gap of 100 bp	
*	28270	29059:	contig of 790 bp	in length
*	29060	29159:	gap of 100 bp	
*	29160	29953:	contig of 794 bp	in length
*	29954	30053:	gap of 100 bp	
*	30054	30834:	contig of 781 bp	in length
*	30835	30934:	gap of 100 bp	
*	30935	31716:	contig of 782 bp	in length
*	31717	31816:	gap of 100 bp	
*	31817	32590:	contig of 774 bp	in length
*	32591	32690:	gap of 100 bp	
*	32691	33471:	contig of 781 bp	in length
*	33472	33571:	gap of 100 bp	
*	33572	34334:	contig of 763 bp	in length
*	34335	34434:	gap of 100 bp	
*	34435	35215:	contig of 781 bp	in length
*	35216	35315:	gap of 100 bp	
*	35316	36083:	contig of 768 bp	in length
*	36084	36183:	gap of 100 bp	
*	36184	36562:	contig of 779 bp	in length
*	36563	37062:	gap of 100 bp	
*	37063	37806:	contig of 744 bp	in length
*	37807	37906:	gap of 100 bp	
*	37907	38682:	contig of 776 bp	in length
*	38683	38782:	gap of 100 bp	
*	38783	39570:	contig of 788 bp	in length
*	39571	39670:	gap of 100 bp	
*	39671	40463:	contig of 793 bp	in length
*	40464	40563:	gap of 100 bp	
*	40564	41353:	contig of 790 bp	in length
*	41354	41453:	gap of 100 bp	
*	41454	42216:	contig of 763 bp	in length
*	42217	42316:	gap of 100 bp	
*	42317	43093:	contig of 777 bp	in length

*	43094	43193:	gap of	100 bp
*	43194	43966:	contig of 773 bp	in length
*	43967	44066:	gap of	100 bp
*	44067	44841:	contig of 775 bp	in length
*	44842	44941:	gap of	100 bp
*	44942	45707:	contig of 766 bp	in length
*	45708	45807:	gap of	100 bp
*	45808	46598:	contig of 791 bp	in length
*	46599	46698:	gap of	100 bp
*	46699	47457:	contig of 759 bp	in length
*	47458	47557:	gap of	100 bp
*	47558	48363:	contig of 806 bp	in length
*	48364	48463:	gap of	100 bp
*	48464	49263:	contig of 800 bp	in length
*	49264	49363:	gap of	100 bp
*	49364	50154:	contig of 791 bp	in length
*	50155	50284:	gap of	100 bp
*	50255	51047:	contig of 793 bp	in length
*	51048	51147:	gap of	100 bp
*	51148	51939:	contig of 792 bp	in length
*	51940	52039:	gap of	100 bp
*	52040	52823:	contig of 784 bp	in length
*	52824	52923:	gap of	100 bp
*	52924	53714:	contig of 791 bp	in length
*	53715	53814:	gap of	100 bp
*	53815	54604:	contig of 790 bp	in length
*	54605	54704:	gap of	100 bp
*	54705	55491:	contig of 787 bp	in length
*	55492	55591:	gap of	100 bp
*	55592	56385:	contig of 794 bp	in length
*	56386	56485:	gap of	100 bp
*	56486	57264:	contig of 779 bp	in length
*	57265	57364:	gap of	100 bp
*	57365	58144:	contig of 780 bp	in length
*	58145	58244:	gap of	100 bp
*	58245	59001:	contig of 757 bp	in length
*	59002	59101:	gap of	100 bp
*	59102	59875:	contig of 774 bp	in length
*	59876	59975:	gap of	100 bp
*	59976	60753:	contig of 778 bp	in length
*	60754	60853:	gap of	100 bp
*	60854	61636:	contig of 783 bp	in length
*	61637	61736:	gap of	100 bp
*	61737	62508:	contig of 772 bp	in length
*	62509	62608:	gap of	100 bp
*	62609	63382:	contig of 774 bp	in length
*	63383	63482:	gap of	100 bp
*	63483	64260:	contig of 778 bp	in length

Query Match	30.8%	Score 630.2;	DB 63;	Length 76653;
Best Local Similarity	99.4%	Pred. No. 1.4e-147;		
Matches 632:	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

[illegible]

QY 1426 accctctgaggatgggctaggaccatgaagcagcaattacaggagagcagacacatag 1485
 |||||
 Db 30424 ACCCTCTGAGGATGGGCTAGGACCATGAAGCAGCAATTCGGAGAGCAGACCTAG 30365
 |||||
 QY 1486 tatacatcttcagatccctgactgagcactaaatccgacacttacatttattt 1545
 |||||
 Db 30364 TATACATTTTCAGATCCCTGCNCTTGGCAGCAATTCACATTTTACATTTT 30305
 |||||
 QY 1546 ttctgtataaataagactactattattaggtataatgaggttttgcatttatattct 1605
 |||||
 Db 30304 TTTCTGTAATTAAGTACTATTAGTAAATAGAGGTTTGTATGCTATTATATATCT 30245
 |||||
 QY 1606 accatcttgaggtagttttaccctgatatagccagaaaaatctagacattctctatct 1665
 |||||
 Db 30244 ACCATCTTGAAGGTAGGTTTACCTGATAGCCAGAAAATATCTAGACATTTCTATATC 30185
 |||||
 QY 1666 attcaggtataatctcttataaacacactattgtttttctataagccatatttttgagca 1725
 |||||
 Db 30184 ATTCAAGTAAATCTCTTTAAACACCTATTGTTTTTCTATAAGCCATATTTTGGAGCA 30125
 |||||
 QY 1726 ctaagtaaaatggcaattgggacagatatattgagg 1761
 |||||
 Db 30124 CTAAGTAAATGGCAATTTGGGACACATAGTGGG 30089
 |||||

RESULT 4

HSLYACLY 2481 bp mRNA PRI 17-FEB-1997
 LOCUS H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
 DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
 ACCESSION Z31690
 VERSION 231690.1 GI:506430
 KEYWORDS LAL; lipase; lysosomal acid lipase.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2481)
 AUTHORS Du, H. and Gregory, G.A.
 TITLE Structural Conservation of Putative Functional Motifs between Mouse
 and Human Lysosomal Acid Lipase
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2481)
 AUTHORS Du, H.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
 Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
 OH, 45229-3039, USA

FEATURES

source
 1. 2481
 /organism="Homo sapiens"
 /strain="HepG2"
 /db_xref="taxon:9606"
 /clone="clone hUAL"
 /tissue_type="liver"
 /cell_type="Hepatoma"
 /clone_lib="human liver HepG2 cDNA library/lambda zap"
 34. .96
 /gene="LAL"
 34. .1233
 /gene="LAL"
 34. .1233
 /gene="LAL"
 /citation=[1]
 /codon_start=1
 /product="lysosomal acid lipase"
 /protein_id="CA83495.1"
 /db_xref="GI:506431"
 /db_xref="SPREMBL:Q16529"
 /translation="MKRFLGLVYCLVLTLSHSGRGLKLTAVDPETNMNVSEIISY
 GPPSEYLVETEDGYILCLNRPVGRKNSDKPKVPVFLQHLGLADSSNWTNLAN
 SLGFIADAGFDVWGMNSRNTWVRKHKTLVSODEFAFSYDEMAKYDLPASINPIL
 NKTGQEQVYVGHSGCTTIGTAFSPQIPELAKRIKMFALGPVAVAFCTSPAKLGR
 LPHLINKDLPQKREFLPQSAFLKWLGHVCTHVLKELCGNLCFLCGLGNERNLNMSR

sig_peptide

gene

CDS

variation

79
 /gene="LAL"
 /citation=[1]
 /replace="c"
 97. .1230
 /gene="LAL"
 /product="lysosomal acid lipase"
 100

mat_peptide

variation
 /gene="LAL"
 /citation=[1]
 /replace="g"
 2459. .2464
 polyA_site
 2480

variation

BASE COUNT 692 a 477 c 513 g 799 t
 ORIGIN

Query Match

25.8%; Score 527.8; DB 93; Length 2481;
 Best Local Similarity 65.4%; Pred. No. 5.3e-122;
 Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1;

QY 164 cagaatggaaatggctctctgattctgtggtggcgatattatccagagaaatgtgaattc 223
 |||||
 Db 30 CAGATGAAAATCGGTTCTTGGGTTGGTGTCTGTGTCTCTGGACCCCTGCATTC 89
 |||||
 QY 224 ag-----tacatatccaaactaaagctgtggaccagaaagcattcatatattagta 277
 |||||
 Db 90 TGAGGGGTCTAGAGGGAACACTGACAGCTGTGGATCTCTGAACAACAAATGATGTAGTGA 149
 |||||
 QY 278 aatcatccaacatcaaggctatccctgtgaggaatatgaagtcgcaactgaagatggta 337
 |||||
 Db 150 AATTATCTTACTTGGGGATTCCTAGTAGGAATACCTAGTTGAGACAGAAAGATGNTA 209
 |||||
 QY 338 tatcttctgtttaacagagattctctcgagcctagtgcaacctaaagagacaggtccag 397
 |||||
 Db 210 TATTCTGTGCTTTAACCGAATTCCTCATGGGAGGAAGAACCATTCCTGACAAAGTCCCAA 269
 |||||
 QY 398 gctgtgggttactcagcagatggcctagttggaggtgtcagaaactggatttccaaact 457
 |||||
 Db 270 ACCAGTTGCTTCTCTCAACATGGCTTGTCTGGCAGATTCCTAGTAACCTGGGTCAAAACCT 329
 |||||
 QY 458 gcccaacaatagcctgggtctcatctctgcagatgcctggtttgacgtgtgagtgaggaa 517
 |||||
 Db 330 TGCCAACAGCAGCCTGGGCTTCATTCTTCTGATGCTGTTTGTACGTGTGGATGGCCAA 389
 |||||
 QY 518 cagcaggggaacgcctggtctcgaacacacagacactctccatagaccagaatgagtt 577
 |||||
 Db 390 CAGCAGAGGAATAACCTGGTCTCGGAACACATAAGACACTCTCAGTTTCTCAGGATGAAT 449
 |||||
 QY 578 ctgggtttcagttatgatgatgagtgctaggttttgacctctcctcagtgatgaactttat 637
 |||||
 Db 450 CTGGGCTTTTCAGTTATGATGAGATGGCAAAATATGACCTACCAAGCTTCCATTAACCTAT 509
 |||||
 QY 638 ttgcagaaaaacggcgcaggaagaagatctattatgcgtattcagcagggcaccacat 697
 |||||
 Db 510 TCTGATATAAACTGGCCAGAACAAAGTGTTATTATGTGGGTCTCTCTCAGGCACCATAT 569
 |||||
 QY 698 gggctttattgcatatttccaccatgccagagctgggtcagaaaaatcaaaatgtatttgc 757
 |||||
 Db 570 AGGTTTATAGCATTTTACAGATCCCTGAGTGGCTAAAGGATTAATAATGTTTTTTC 629
 |||||
 QY 758 tttagcaccatagcactgttaagcatgcaaaaagcccccggaccacaaatatttctgtct 817
 |||||
 Db 630 CCTGSGCTCTGTGGCTTCCGCTTCTGTACTAGCTATGGCCAAATATAGACGATT 689
 |||||
 QY 818 gccagatgatgatcaaggatttttggcaaaaaagaatttctgtatcagaccagatt 877
 |||||
 Db 690 ACCAGATCATCTATTAAAGACATTATTGGAGACAAAGAAATTTCTTCCCCAGAGTGGCTT 749
 |||||
 QY 878 tctcagacaaactgttatttaccatttgggtggcaggtgattcttgatcatcagattttagtaa 937
 |||||


```

Db 742 CTGGCTCTGGCTTCCTCGCTTCTGTCTAGCCCTTAGCCCAATTAGACGATT 801
Qy 818 gccagatgatgatcaaggagattgttggcaaaaagaatttctgtatcagaccagatt 877
Db 802 ACCAGATCATCTAATAGGACTTATTTGGAGACAAGAATTTCTTCCAGAGTGGCTT 861
Qy 878 ttcagacaacttatttacccttctgtgcccaggtgatttcttgcagattttagtaa 937
Db 862 TTTGAAGTGGCTGGGTACCCACGTTTGCACATCTGTCTATCTGAAGGAGCTCTGTGAAA 921
Qy 938 tatcatgttacttcttgggtgattcaacacacaaatgaacatgaacgagccgacaagt 997
Db 922 TCTCTGTTTTCTTGTGTGGATTATATGAGAGAAATTTAAATATGTCTAGAGTGGATG 981
Qy 998 atatgtgtccacactcttctgtggaacatctgtgcaaaatattctacactggagccaggc 1057
Db 982 ATATACAACACATCTCTCTGCTGGAACCTTGTGCAAAACATGTATACACTGGAGCCAGGC 1041
Qy 1058 agtgaattctgtgaactccgggcatcttgactggggggagtgagacaaaatctggaaaa 1117
Db 1042 TGTAAATTCCAAAGTTTCAAGCCTTTGACTGGGGAAGCAGTCCCAAGAATTTATTTCA 1101
Qy 1118 atgcaatcagccaaactctgtaaggtagagtcagagatgacggtccctcacagcaat 1177
Db 1102 TTACACACAGAGTATCTCCACATACATATGTGAAGGACATGCTTGGCCGACTGCAGT 1161
Qy 1178 gtgacaggaggtcaggactggcttcaaatccagaagacgtgaaaatgctgctctctga 1237
Db 1162 CTGGAGCGGGGTACAGACTGGCTTGCAGATGTCTACAGCTCAATATCTTACTGACTCA 1221
Qy 1238 ggtgaccaactcattaccatagaatattcttgaatgggtcagctggatttcattctg 1297
Db 1222 GATCACCACACTGGTGTTCATGATGAGCATTCCGGAATGGGAGCATCTTGACTTCAATTG 1281
Qy 1298 ggggttgatctctcaccatgatgaatgaatgaatcattccatctgatgcagcaggaga 1357
Db 1282 GGGCTGTGATGCCCTTGGAGCTTTATATATAAATTTAATCTATGAGGAATATCA 1341
Qy 1358 gaccaacct 1366
Db 1342 GTGAAAGCT 1350

RESULT 7
LOCUS HUMLIPCHL 2493 bp mRNA PRI 07-JAN-1995
DEFINITION Human lysosomal acid lipase/cholesterol esterase mRNA, complete cds
ACCESSION M74775
VERSION M74775.1 GI:187151
KEYWORDS lysosomal acid lipase/cholesterol esterase.
SOURCE Homo sapiens fibroblast cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2493)
AUTHORS Anderson,R.A. and Sando,G.N.
TITLE Cloning and expression of cDNA encoding human lysosomal acid
lipase/cholesterol ester hydrolase. Similarities to gastric and
lingual lipases
JOURNAL J. Biol. Chem. 266 (33), 22479-22484 (1991)
MEDLINE 92042192
FEATURES
    Location/Qualifiers
        1..2493
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="fibroblast"
            /map="10"
            <1..2493
            /gene="LIPA"
            /note="G00-120-153"
            1..2493
    mRNA
    gene

```

```

CDS
/gene="LIPA"
41..1240
/EC_number="3.1.1.13"
/notes="alternate initiator ATG may be at position 47"
/codon_start=1
/db_xref="GDB:G00-120-153"
/product="lysosomal acid lipase/cholesterol esterase"
/protein_id="AAA59519.1"
/db_xref="GI:187152"
/translation="MKMRFLVVLVPLHSEGGSGKLTAVDPETNMVSEIISYW
GFSEYLVETEDGYILCLNRIPIGRKNSDKGPKVVFQHGGLADSSNMVETNLANS
SLGFIADAGFDVWMSNRGNTWSRKHKHTLSVODEFWAFSDYEMAYIDLPASINFIL
NKTGQSOVYIVGHSGQGTIGFIAPSQIPELAKRIKMFALFASVAFSPMAKLGR
LPDLHLKDLFGDFEFLPOSQFLKWLGHVTHVILKELCGNLCLFLGCFERNLNMSR
VDYTTTHSPAGTSQVQNLHWSQAVKFOAFDWGSSAKNYFHYNQSPPTYNVKDML
VPTAVWSGGHDLADVDVNNILLITQINLVFHESHIPWEHLDFIWLGLDAPWRLYNKII
NLMRKQY"
41..103
sig_peptide
/gene="LIPA"
/notes="alternate signal peptide may begin at position 47;
as predicted from von Heijne, G. (1986) Nucl. Acids Res.
14, 4683-4690; G00-120-153"
104..1237
mat_peptide
/gene="LIPA"
/EC_number="3.1.1.13"
/notes="G00-120-153"
/product="lysosomal acid lipase/cholesterol esterase"
2466..2471
polyA_signal
/gene="LIPA"
/notes="G00-120-153"
2493
polyA_site
/gene="LIPA"
/notes="G00-120-153"
BASE COUNT 692 a 484 c 516 g 801 t
ORIGIN

Query Match 25.7%; Score 526.2; DB 97; Length 2493;
Best Local Similarity 65.3%; Pred. No. 1,3e-121;
Matches 790; Conservative 0; Mismatches 413; Indels 6; Gaps 1;

Qy 164 cagaatgaaatgtggtcttctgattctgtggtggttatgttccagagaaatgtgaattc 223
Db 37 CAGAAATGAAATGCGGTTCTTGGGTTTGGTGTCTGTGTGTCTCTCGCCCTCATTC 96
Qy 224 ag-----tacatatgcaactaaagctgtgacccagaaagcattcatgaattatgga 277
Db 97 TGAGGGGTCTGGAGGGGAAACTGACAGCTGTGTGATCCTGAAACAAACATGAATGTGAGTGA 156
Qy 278 aatcatccaacatcaaggctatccctgtgaggaatatgaagtcgcaactgaagatgggta 337
Db 157 AATTATCTTACTTGGGGATTCCTTAGTGAGGAATACCTAGTTGAGACAGAGATGGATA 216
Qy 338 tctcttctgttaacaggatctctcagggcctagtcgcaactgaagacagagttccag 397
Db 217 TATTCTGTGCTTAACCGAATTCCTCATGGGAGGAAGAACCATTCGTGACAAAGTCCCAA 276
Qy 398 gctgtggtgttaactgcagatgacctagtgtggaggtgctagcaactgattccaaact 457
Db 277 ACCAGTTGTCTTCTGCAACATGGCTTCTGCTGCGAGATTCCTAGTAACCTGGGTCACAAAGCT 336
Qy 458 gcccacaacatagcctgggttcattcttggcagatgctggttttgagctgtgagtgga 517
Db 337 TGCCACACAGAGCCTGGGCTTCATCTTCTGTGATGCTGGTTTACGCTGTGGATGGGAA 396
Qy 518 cagcagggaacgcctgtgtctcgaaaacacaaagacactctccatagaccagatgagtt 577
Db 397 CAGCAGAGGAATACCTGTGCTCGGAACATAAGACATCTCTCAGTTTCTCAGGATGATT 456
Qy 578 ctgggcttcagttatgatgatgagtggtcaggttttgaccttctcgtcagtgataaactttat 637
Db 457 CTGGGCTTTCAGTTATGATGATGGCAAAATATGACCTACACCTACAGCTTCCATTAACTTCAT 516

```

```
Qy 638 tttcagaaaaacggccaggaagaaagatctattatgtcgggtattcagcaggccaccacat 697
Db 517 TCTGAATAAAATGCGCAAGAAAGTATATATGTGGTTCATCTCAAGCCACACTAT 576
Qy 698 gggctttattgcatcttccaccatgcccagagctggctcagaaaaatcaaaatgtatttgc 757
Db 577 AGGTTTATACATTTTCACAGATCCCTGAGCTGGCTTAAAGGATTAAATGTTTTTTCG 636
Qy 758 tttagcaccatagaccactgttaagcatgcagaaagcccccggggaccacaaatctttgttgc 817
Db 637 CTTGGTCTCTGCTCCGTCGCTTCGTACTAGCCCTATGCCCCAAATTAGACCAATT 696
Qy 818 gccagatatgatcatcaaggattgtttggcagaaaaaagaatttctgtatcagaccagatt 877
Db 697 ACCAGATCATCTCATTAAGACATTATTGGAGACAAAGAAATTTCTTCCCCAGAGTGGCT 756
Qy 878 tctcagacacactgttattacccttggccaggtgattcttgatcagattttagtaa 937
Db 757 TTTGAAGTGGTGGGTACCCAGCTTTGCACTCATGTCTACTGAAGGAGCTCTGTGGAAA 816
Qy 938 tatcatgttactcttgggtgattcaacacacacaaatgatgaacatgagccgagcaagtgt 997
Db 817 TCTCTGTTTCTCTGTGTGGATTATAGAGAGAAATTTAAATATGTCTAGAGTGGATGT 876
Qy 998 atatgctgccacacactcttctggaacatctgtgcaaaaattattacactggagccagggc 1057
Db 877 ATATACAAACACATTTCTCTGCTGGAACTTCTGTGCAAAACATGTTACACTGGAGCCAGGC 936
Qy 1058 agtgaattctgtgaactcgggcatcttgactggggagtgagaccacaaatctcgaaaa 1117
Db 937 TGTAAATTCCAAAGTTTCAAGCTTTGACTGGGGAGGAGTCCCAAGAAATATTATTCA 996
Qy 1118 atgcaatcagccaactctgaagtacagagtcagagatatgacggtccctacagcaat 1177
Db 997 TTACAACACAGATTATCTCCACATACAAATGTGAAGGACATGCTTGTCCGAGTGCAGT 1056
Qy 1178 gtgacagagggtcagagactggtttcaaatccagagaagcgtgaaatgtgctctctga 1237
Db 1057 CTGGAGGGGGGTCCAGCTGGCTTTCAGATGTCTACGACGTCAATATCTTACTGACTCA 1116
Qy 1238 ggtgaccaacctcctaccataagaattcctggaatggctcagctgatttcatctg 1297
Db 1117 GATCACCAACTTGTGTTCATGAGACATTCGGAAATGGAGGATCTTCACTTCATTG 1176
Qy 1298 gggtttgatgctcctcacogtatgtacaatgaatcatccatcgtatgagcaggagga 1357
Db 1177 GGGCTGGATGCCCTTGGAGGCTTTATAATAAAATTTAATCTAATGAGGAATATCA 1236
Qy 1358 gaccacact 1366
Db 1237 GTGAAGCT 1245

RESULT 8
A26689
LOCUS A26689 1378 bp DNA PAT 05-APR-1995
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE Benicourt,C., Blanchard,C. and Junien,J.L.
AUTHORS Recombinant gastric lipase from rabbit and pharmaceutical
TITLE compositions
JOURNAL Patent: EP 0542629-A 9 19-MAY-1993;
INSTITUT DE RECHERCHE JOUVEINAL
FEATURES Location/Qualifiers
source 1..1378
/organism="synthetic construct"
```

```
BASE COUNT 397 a 287 c 276 g 418 t
ORIGIN
Query Match 23.9%; Score 489.2; DB 9; Length 1378;
Best Local Similarity 65.0%; Pred. No. 2.5e-112;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
Qy 250 acccagaagcattcatgaataattagtgaaatcatcccaacatcaagctatccctgtgaag 309
Db 136 ACCCTGAAGTGATATATATATAGTCAGATGATTTCTTACTGGGGATACCCAAGTGA 195
Qy 310 aatatgaagtcgcaactgaagatgggtatatctcttctgttaacaggaattctctcgagcc 369
Db 196 AATATGAAGTTGTGACTGAAGATGGTTATATCTCTTGAAGTTAACAAGATTCCTTATGGGA 255
Qy 370 tagtgcacacctagaagacagggttccaggcctgtggtgttactcagcagatggcctagt 429
Db 256 AGAAAAATTCAGGGGAACAGAGGCCAGAGACCGCTCGTATTTTTCAGCATGGTTTGTG 315
Qy 430 gaggtgctagcaactgatttccaaacctgcccacaaatagcctgggtcttattctgcag 489
Db 316 CATCAGGCTCAAACTGATCTCCAACTGCCCAACAACAGCCCTGGCTTTATTCTGGCAG 375
Qy 490 atgctgggttttgactgctggtgaggggaacagcaggggggaacgcctgctctgaaaaaca 549
Db 376 ATGTGGTTATGTGTGTGGTGGGAACAGCAGAGGAATACTGTGTCCAGAAGAACT 435
Qy 550 agacactctccatagaccagaagatgagttctgggcttccagttcagttatgatgagatgg 609
Db 436 TATACTATTACCAGACTCCGTTGAATTCGGGCTTTTAGCTTTTGATGAATGGCTAAAT 495
Qy 610 ttgcctctctcagtgataaactttatttgagaaaacggcggcagaaaagatcatt 669
Db 496 ATGACCTTCCAGCCACAAATTCAGCTTCATTGTAAGGAAACTGGACAGAAAGCTGCACT 555
Qy 670 atgtcggctattcacaggggcaccaccatgggctttattgcaattttccaccatgcccagc 729
Db 556 ATGTGTGTCATCTCCAGGGCACCACCATTGTTTATTTATTTGCTTTCTACAAATCCCAAGC 615
Qy 730 tggctcagaaaaacaaatgtattttgctttagcaccatagccactgttaagacatcaa 789
Db 616 TGCTGAAAGAAATCAAAACCTTTTATGCATTAGCTCCAGTCGCCACTGTGAAGTATACA 675
Qy 790 aaagcccgaggacaaatcttctgctgcagatatgatcaagggatgttttgca 849
Db 676 AAGCCCTTGTAAACAAACTTAGGTTTATCTCTCCCAACCAATGTTCAAGATTATTTGGTG 735
Qy 850 aaaaagaattctgtatcagaccagatttctcagacacactgttatttaccctttgtgccc 909
Db 736 ACAAAATATTTCTACCCACACAAATTTCTTGATCAATTTCTGCCACTCAAGTGTGTTCCC 795
Qy 910 aggtgattctgtatcagattgttagtaataatcatctgttacttctctgggtggattcaacca 969
Db 796 GTGAGACACTGAATGTAAATTTGCAGCAATGCCCTTATTTATCATTTGTGGCTTTCCACAGC 855
Qy 970 acaatatgaacatgagccgagcaagtgtatatgctgccacactcttctgtggaacatctg 1029
Db 856 CAAACTTGAACATGAGTCGCTGGATGCTGACGTGTACGTGTACATATCCCGCAGGAACACTCAG 915
Qy 1030 tgcataatattctacatggagccaggtgagtaattctggtgaactccgggactttagt 1089
Db 916 TTCAAAACATGCTGCACACTGGACCCAGGCTGTTAAATCTGGAATTTTCAAGCTTTTAAT 975
Qy 1090 ggggagtgagacaaaaatctgagaaaaatgcaatcagcacaactcctgtgaaggtacagag 1149
Db 976 GGGGAGTCCAGCTCAGAAATGTAGTGCACTTCAATAGCCACACACTCCCTACTACCAATG 1035
Qy 1150 tcagagatatgacggttccttacagcagatgtggacagaggtcagaggtcggctttcaaatc 1209
Db 1036 TGACCCCATGATGTGCCAATTTGCAGTGTGGAGTGTGGCAATGACTGGTGTGCTGACC 1095
```



```
/protein_id="CAA29413.1"
/db_xref="GI:758063"
/translation="MWLLTMSLISVLGTHGLFGLKHPGSPVETMNISSOMITYWGY
PNEEYVVDYDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLS
AFTLDAGYDVLGNSRGTWARNNLYSPDSVEFWAFSDFEMAKYDLDPATIDFIVK
TGOKOLHYVGHSGOGTIGFAESTNPSLAKRIKTFYALAPVATYKTSKLNKLFV
QSLFKIFGDKIFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSKNFTSRLD
VYLSPHAGTSONVPHWTOAVKSGFOAYDMGSPVQNMHYDQSPPYNTAMNVP
IYNWNGKDLADPDQVGLLPLKPLNLIYHKEIPYNHLDFIWMADPOEVYNDIVSM
ISDDKK"
63..1241
/codon_start=1
/product="gastric lipase precursor"
/protein_id="CAA29414.1"
/db_xref="GI:758064"
/translation="MASLISVLGTHGLFGLKHPGSPVETMNISSOMITYWGYNBEYE
VYTEDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLSAFIAD
AGYDVLGNSRGTWARNNLYSPDSVEFWAFSDFEMAKYDLDPATIDFIVKTKGQKL
HYVGHSGOGTIGFAESTNPSLAKRIKTFYALAPVATYKTSKLNKLFVQSLFK
IFDKIIFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSKNFTSRLDYSK
PAGTSVQNMHYDQAVKSGFOAYDMGSPVQNMHYDQSPPYNTAMNVPYVING
GKDLADPDQVGLLPLKPLNLIYHKEIPYNHLDFIWMADPOEVYNDIVSMISDDKK"
63..101
/sig_peptide
102..1238
/mat_peptide
396 a 284 c 268 g 417 t
BASE COUNT
ORIGIN
```

CDS

sig_peptide

mat_peptide

BASE COUNT

ORIGIN

```
Query Match 23.0% Score 470; DB 93; Length 1365;
Best Local Similarity 64.2% Pred. No. 1.7e-107;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 244 ctgtggccagagcattcatgaattagtgaaatcatccacatcaaggctatccct 303
DB 121 CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTTGGGGATACCCAA 180

QY 304 gtgggaatgaagtcgaactgaagtggtatctcttctgttaacaggtatctc 363
DB 181 ATGAAGATATGAAGTTGTGACTGAAGATGTTATATCTTGAGTCAATAGAAATTCCTT 240

QY 364 gagcctagtgaacctaaagaacaggttccagggctgtgtgttactcagcagtgcc 423
DB 241 ATGGGAAGAAAAATTCAGGGAATACAGCCAGAGACCTGTGTGTGTTTTCACAGATGGTT 300

QY 424 tagttgagtgctagcaactggaattccaaacctgcccacaaatagcctgggttcattc 483
DB 301 TGCTTGATCAGCCACACAACTGGATTTCCAACTGCCGAACACAGCCTTGCCCTTCATTC 360

QY 484 tggcagatgctgtgttgagtgatggaggaacaggggaaacgctgtgtctgaa 543
DB 361 TGGCAGATGCTGGTATGATGTGGTGGCAACACAGAGAAACACCTGGGCCAGAA 420

QY 544 aacacagacactctccatgacacaaagatgagttctgggtttcagttatgatgagatg 603
DB 421 GAAACTTGACTATTCACAGATTCAGTTGATTCGGCTTTCAGCTTTGATGAATGG 480

QY 604 ctgagtttgacctctctgagtgataaactttatttgcagaaaaagggccaggaagaaga 663
DB 481 CTAATAATGACCTTCCAGGCCACATCGACTTCATTGTGAAGAAAACTGGACAGAGCAGC 600

QY 664 tctattatgctgctattccagggcacccatggcttattgcatcttccaccatgc 723
DB 541 TACACTATGTGGCCATTCACAGGCCACCACTTGTGTTTATGCTTTTTCACCAATC 600

QY 724 cagagctggctcagaaaaatcaaatgattttgtcttagcaccatagcactgttaagc 783
DB 601 CCAGCTGGCTAAAGAAATCAAACTCTATGCTTAGCTCTCTGTGGCCACTGTGAAGT 660
```

```
QY 784 atgcaaaagcccccggggacaaaatttttggctgcccagatatgatgatacaaggattgt 843
DB 661 ATACAAAAGCCCTATATAACAAACAACTAGATTTCCTCAATCCCTCTTCAAGTTATAT 720

QY 844 ttggcaaaagaagattctctgataccacagagattcttcagacacactgttattacatt 903
DB 721 TTGGTGACAAAATATTCTTCCACACAACTTCCTTGATCAATTTCTTGCTGAAGTGT 780

QY 904 gtggcagggtgattcttctgatacagattgttagtaatatcatgttactcttgggtgattca 963
DB 781 GCTCCCGTGAGATGCTGAATCTCTTTGGAGCAATGCCCTATTATTAATTTGGATTG 840

QY 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtctgccacactcttgcgtgaa 1023
DB 841 ACAAGTAAGAACTTTAAACACAGTCGCTGGATGTGTATCTATCATACATAAATCCACAGGAA 900

QY 1024 catctgtgcaaaatattctacactgagccagcagcagtgaaattctgtgtaactccgggcat 1083
DB 901 CTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTAAAGTCTGGGAAATTCACAGCTT 960

QY 1084 ttgactgggggagtgagaccacaaatctggaataatgcaatcagccaaactcctgtaaggt 1143
DB 961 ATGACTGGGGAAGCCAGTTTCAAGTAGGATGACTATGATGATGATGATGATGATGATGATGAT 1020

QY 1144 acagatcagagatatgacggttccctacagcaatgtggacagaggtcaggaactggctt 1203
DB 1021 ACANTGTGACAGCCATGAATGTACCAATTCAGTGTGGAACGGTGGCAAGGACCTGTGTGG 1080

QY 1204 caaatccagaaacgtgaaatctctctctgagtgagcgaacccatcattaccataaga 1263
DB 1081 CTGACCCCAAGATGTGGCTTTTCTCCAAACATCCCAATCTTATTTTACCAAGG 1140

QY 1264 atattctgaatgggtcactggtgattctctgggtttgggtgctccctcaccgtatgt 1323
DB 1141 AGATTCTTTTACATCACTTGTGACTTTATCTGGCAATGGATGCCCTTCAGAACTTT 1200

QY 1324 acaatgaatcatccatctgat 1345
DB 1201 ACAATGACATTTGTTCTATGAT 1222
```

RESULT 11

A01046

LOCUS

DEFINITION

AC1046

H.sapiens mRNA for human gastric lipase.

AC1046

AC1046.1 GI:344241

gastric lipase.

human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Location/Qualifiers

1..1367

/organism="Homo sapiens"

/db_xref="taxon:9606"

47..103

/note="gastric lipase"

/product="lipase"

/protein_id="CAA00125.1"

/db_xref="GI:344242"

/db_xref="SWISS-PROT:P07098"

/translation="MWLLTMSLISVLGTHGLFGLKHPGSPVETMNISSOMITYWGY

PNEEYVVDYDGYILEVNRIPYKKNNGTGORPVVFLQGLLASATNWSLNPNNLS

AFTLDAGYDVLGNSRGTWARNNLYSPDSVEFWAFSDFEMAKYDLDPATIDFIVK

TGOKOLHYVGHSGOGTIGFAESTNPSLAKRIKTFYALAPVATYKTSKLNKLFV

QSLFKIFGDKIFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSKNFTSRLD

ISDDKK"

sig_peptide

CDS

VYLSHPAGTSVONNHFHTQAVKSKFQAYDWGSPVQNMHYDQSPYYNVTAMNVP
IAVWNGKDLADPODVGLLPLNLIYHKEIPFYNHLDFIWMADPAQOEYNDIVSM
ISEDKK"

mat_peptide

104.1240
/note="gastric lipase"
/product="lipase"

BASE COUNT 397 a 284 c 269 g 417 t

ORIGIN

Query Match 23.0%; Score 470; DB 9; Length 1367;

Best Local Similarity 64.2%; Pred. No. 1.7e-107;

Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 244 ctgtgaccagaagcattcgaatatttagtgaataatcccaacatcaaggctatccct 303
DB 123 CTGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTATTGGGGATACCCAA 182
QY 304 gtgaggaataatgaagtcgaactgaagatgggtatatactcttctgttaacaggattccctc 363
DB 183 ATGAAGAATTAAGATGTGTGACTGAAGATGGTTATATCTTGAAGTCAATAGAAATTCCTT 242
QY 364 gaggcctagtcgaactaagaacagaggtccagcctgtggtgttactgcagcatggcc 423
DB 243 ATGGGAGAAAATTCAGGGAATACAGGCAGAGACCTGTGTGTTTTTGCACATGGTT 302
QY 424 tagttggaggtctagcaactgatttccaaacctgcccacaataatagcctgggtctcattc 483
DB 303 TGCTTGATCAGCACCAAACTGGATTCCAACTGCGGAACACAGACCTGCTTCATTC 362
QY 484 tggcagatgctgttttgacgttgatggggaacagcggggaacgcctgtgctcga 543
DB 363 TGGCAGATGCTGTGTATGATGTGTGCTGGGCAACAGACAGAGAAAACACCTGGGCCAGAA 422
QY 544 aacacaagacactctccatagacacagatgagttctgggcttccagtattgatgatgagtg 603
DB 423 GAACTTGATCTATTCACAGATTCAGTTGAATCTGGGCTTTCAGCTTTGATGAATGG 482
QY 604 ctgagtttgacctctctgcagtgatgaacttattttgcagaaaacgggcccaggaaga 663
DB 483 CTAATATGATGCTTCCAGGCACCAATCGACTTCATTGTGAAGAAAACCTGGACAGAGCAGC 542
QY 664 tctattatgctgcgtattccacagggcaccacacacacacacacacacacacacacacacac 723
DB 543 TACATATGTTGGCCATTCACAGGCACACACACATGGTTTATGTGCTTTTCCACCAATC 602
QY 724 cagagctggctcgaataacaaatgtatttctttagcaccatagcactgttaagc 783
DB 603 CCAGCTGGCTAAAGATCAAACTCTATGCTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 682
QY 784 atgcaaaaagccccgggaccacaaatatttggctgcccagatatgatgatacaaggattgt 843
DB 663 ATACAAAAGCCCTATATAACAACTAGATTGTTCTCTCAATCCCTCTTCAAGTTTATAT 722
QY 844 ttggcaaaaagatttctgtatcaagcagatttctcagacaaactgttattaccctt 903
DB 723 TTGGTGCAAAAATATTCTACCCACACAACTTCTTTGATCAATTTCTTGCTACTGAAAGTGT 782
QY 904 gtggcaggtgattcttgatcagattgttagtaataatcatgttactcttgggtggattca 963
DB 783 GCTCCCTGAGATGCTGAATCTCCTTTGACGCAATGCTTATTATTAATTTGTGGATTG 842
QY 964 acaccaacaatatgaactgagccgagcaagtgatatgctgcccacactcttgcgtggaa 1023
DB 843 ACAGTAAGAATTTAACACAGCTGCTTGGATGTGTATCTATCACATAATCCAGCAGGAA 902
QY 1024 catctgtgcaaaaattctctacactgagccagcagtgaaattctgtgtaactccgggcat 1083
DB 903 CTTCTGTGTAACAACTGTTCCATTGGACCCAGGCTGTTAAGTCTGGGAAATTCAGACTT 962
QY 1084 ttgactggggagtgagaccacaaactggaataatgcaaatcagcaactcctgtgaagt 1143
DB 963 ATGACTGGGAGGCCAGTTTCAGAAATAGGATGACATATGATCAGTCCCAACCTCCCTACT 1022

QY 1144 acagagtcagagatatgacgggtccctacagcaatgtggacagaggtcagagctgggttt 1203
DB 1023 ACAATGTGACAGCAGTGAATGACCAATGTGCAAGCGGTGCAAGACCTGTGG 1082
QY 1204 caaatccagaagacgtgaaatctgctctctgagtgagtgacacacctcatatcacaaga 1263
DB 1083 CTGACCCCAAGATGTGGCCTTTGCTTCCAAACATCCCAATCTTATTATACCACAGG 1142
QY 1264 atattctgaatgggtcagtcacgtgatttcaatctgggttggatgctccctcacgtatgt 1323
DB 1143 AGATTCTTTTACATCACTTGGACTTTATCTGGCAATGGATGCCCTCAAGAAGTTT 1202
QY 1324 acaatgaaatcatccatctgat 1345
DB 1203 ACAATGACATTTTCTATGAT 1224

RESULT 12

LOCUS

Al2714 1367 bp DNA PAT 09-AUG-1994
DEFINITION Pregastric lipase.

ACCESSION Al2714

VERSION Al2714.1 GI:579560

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1367)

AUTHORS

TITLE

PRECURSOR POLYPEPTIDE, DNA SEQUENCE CODING THEREFOR, VECTORS, HOST

ORGANISMS, AND PROCESSES INVOLVING SAME

Patent: WO 8603778-A 8 03-JUL-1986.

JOURNAL

Location/Qualifiers

Source

1..1367

/organism="Homo sapiens"

/db_xref="taxon:9606"

47..1243

/codon_start=1

/product="pregastric lipase"

/protein_id="CAA01053.1"

/db_xref="GI:579561"

/db_xref="SWISS-PROT:P07098"

/translation="MWLLTMAISLVLTTHGLFKLHFGSPVMTNISMITYWGY

PNEEVVTEEDGYILEVNRIPYKNSGNTGQRPVVFLOHGLLASATNWIISLPNNSL

ATFLADAGYDVLWNGSKNTWARRNLYSPDSVEFWAFSDENAKYDLPATIDFIVKK

TKOKLHYVGHSGQTTIGPTAEFTNPSLAKRIKTFVALAPVATVKYTKSLINKLRFVP

QSLFRIFGDKIFYPHNFDQFLATEVCSREMLNLCNSNALFICIGDSDRNFTNRLD

YLSLHNPAGTSVONNHFHTQAVKSKFQAYDWGSPVONRMHYDOSQPPYYNVTAMNVP

IAVWNGKDLADPODVGLLPLNLIYHKEIPFYNHLDFIWMADPAQOEYNDIVSM

ISEDKK"

BASE COUNT 397 a 284 c 269 g 417 t

ORIGIN

Query Match 23.0%; Score 470; DB 9; Length 1367;

Best Local Similarity 64.2%; Pred. No. 1.7e-107;

Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 244 ctgtgaccagaagcattcgaatatttagtgaataatcccaacatcaaggctatccct 303

DB 123 CTGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTATTGGGGATACCCAA 182

QY 304 gtgaggaataatgaagtcgaactgaagatgggtatatactcttctgttaacaggattccctc 363

DB 183 ATGAAGAATTAAGATGTGTGACTGAAGATGGTTATATCTTGAAGTCAATAGAAATTCCTT 242

QY 364 gaggcctagtcgaactcaagaagacaggttccagcctgtggtgttactgcagcatggcc 423

DB 243 ATGGGAGAAAATTCAGGGAATACAGGCAGAGACCTGTGTGTTTTTGCACATGGTT 302

QY 424 tagttggaggtctagcaactgatttccaaacctgcccacaataatagcctgggtctcattc 483


```
Qy 844 ttggcaaaaagaattctctgtatccagaccagattctccagacaactgtgttattacacctt 903
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 680 TTGGCAAGAAATGTTCTCTCCACACACCTACTTTGATGACTTCTTGGTACCGAAGTGT 739
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 904 gtggccagggtgattctgtacagattgttagtaatacatctgttactctgtggtgattca 963
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 740 GCTCAGCGGAGGTCTAGATCTCTCTGCGACACACTTTATTCATCTCTGTGGATTG 799
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtgcccacactcttgcggaa 1023
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 800 ACAAGAAAAAATTAAATGTGAGTCGTTTGTATGTATCTAGGCGATAATCCAGCAGGA 859
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1024 catctgtcaaaatattctcaactgagccagcagtggaattctgtgtaactccgggcat 1083
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 860 CATCTGTTCAAGACTTCTCCACTGGGCACAGCTTGTAGATCTGGGAAATTTCAAGCCT 919
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1084 ttgactggggagtgagaccacaaaatctggaataatgcaatcagccactcttgaaggt 1143
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 920 TCAACTGGGAGGCCATCCAGACATGTTACACTACACACAGAAACGCTCCCTGAAT 979
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1144 acagagtcaagatatgacgtgctccctacagcaatgttgagcagaggtcaggactggctt 1203
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 980 ATGATGTGTGACCATGACTGTGCCAGTTGCGAGTGTGGAACGGTGGCAATGACATCCTG 1039
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1204 caaatccagaagaagctgaaatgtctgtcgtgaggtgagcaacacctcatciaccataaga 1263
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1040 CTGATCCCCAAGATGTGCGCATGTGCTTCCCAAACTCTCCAACTCTCTTCCATTAAG 1099
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1264 atattctgaatgggctcagtgattctcctgggttgggtctcctccacogtatgt 1323
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 AGATCTTGCTCAATCACTGACTGACTTCATCTGGGCAATGGATGCCCTCAAGAGTTT 1159
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1324 acaatgaatcatccatctgtatgagcagcagcagcagcagcagcagcagcagcagc 1365
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1160 ACAATGATGATTTCCATGATGCGAGAAAGACTAAAGAATC.1201
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
RNLIP RNLIP 1355 bp mRNA ROD 30-MAR-1995
LOCUS Rat mRNA for lingual lipase.
DEFINITION X02309
ACCESSION X02309
VERSION X02309.1 GI:56595
KEYWORDS glycoprotein; lipase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1355)
Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C.,
Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J.
Molecular cloning and nucleotide sequence of rat lingual lipase
cDNA
JOURNAL Nucleic Acids Res. 13 (6), 1891-1903 (1985)
MEDLINE 85215387
COMMENT Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.
FEATURES
source 1..1355
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
sig_peptide 49..102
CDS 49..1236
/product="put. signal peptide (aa -18 to -1)"
/db_xref="lipase precursor"
/codon_start=1
/protein_id="CA26179.1"
/db_xref="GI:56596"
/translaton="MWLLLTIVSTFGAGHLGKLPNGNPEANMNISOMITYWGY
COEVEVTEGYILGVRIPIPHKNSENIGKRPVYLOHGLIASATNWTANLNNLSIA
FMLADAGYDWLNSRGNTWSRKNVYSPDSVEFASFDEMAKIDLPATINFIQKT
GQEKIHVHVSQGTITGFIATFNPTLAKKIKTFYALAPVATVKYITQSPLUKISFPT
```

```
FLKLMFGKKMFLPHTYFDPLGTEVCSEVRLDLQSNLFLIFCGDDKKNLNVSRFDV
YLGHPAGTSVQDFLHWAOLVRSKGFQAFNWSGSPQSNMLHYNOKTPEYDVSAMTVPV
AYVNGGNDILADPDQVAMLLPKLSNLLFKELILAYNHLDFINWADAPQEVYINEMISM
AED"
mat_peptide 103..1233
/product="lipase (aa 1-377)"
misc_feature 505..507
/note="aa 153 pot. lipid binding site"
misc_feature 1355
/note="polyA site"
BASE COUNT 400 a 291 c 387 t 1 others
ORIGIN
Query Match 22.9%; Score 467.6; DB 95; Length 1355;
Best Local Similarity 63.5%; Pred. NO. 6.8e-107;
Matches 71;; Conservative 0; Mismatches 409; Indels 0; Gaps 0;
Qy 244 ctgtggaccagagcattcatgaattagtgaaatccatcccaacatcaaggctatccct 303
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 CTGCAAAACCTGAAGCAAAATATGAATATTAGTCAGATTAACCTTACTGGGGATATCCAT 181
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 gtgggaatatgaatgcgaactggaagtgggtatatcccttctgttaacagaggtctc 363
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 GTCFAGANATAGAGTTGTTACTGGAAGATGGCTACATCTCGGGGCTCTACAGAAATCTCTC 241
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 364 gaggcctagtgcacaccttaagaagacaggttccaggcctgtggtgttactcagcagtcgc 423
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 ATGCGAAGAAATAATCTGAAAAATATAGGCAAGACACCTGTGGTGTATTGTCAGCATGGTT 301
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 424 tagttggaggtgctagcaactggtattccaaacctgcccacaaatagctgggtctcattc 483
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 TGATCGCATCAGCCACAAACTGGATTGCAAAATCTTACCAAAACACAGCCTGGCCCTTATGC 361
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 484 tggcagatgctgttttgacgtgtggatggggaacagcagggggaacgcctgggtctcga 543
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TAGCAGACGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGAATACATGTTGCCAGGA 421
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 544 aacacaagacactctccatagaccacagatgattcttgggtcttcagttatgatgagatgg 603
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AAAATGTATACTACTACCAGACTCAGTTGAATTTCTGGCTTTCAGCTTTTATGAAATGG 481
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 504 ctagggttgacctctctgcagtgataaaacttatttgcagaaaaacgggcaggaagaaaga 663
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 CTAAATATGACCTTCCCGCCACAAATAAATCTTATTGTACAGAAAACCTGGACAAGAGA 541
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 664 tctattatgcgctattctacagggccaccacacacacacacacacacacacacacacac 723
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 TACTACTATGTGTCATCTTCAGGGCACACACTATTGTTGTTTCATTCACCAATC 601
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 724 cagagctggctcagaaaaatcaaaatgtatttggcttttagcaccacacacacacacacac 783
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 CTACACTGGCCAAAAAATCAAGACGTTTTATGATCATAGCTCCAGTTCTACCGTGAAGT 661
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 784 atgcaaaagccccgggagcaaaatcttctgtcgcagacatagatgaagggattgt 843
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATACACAAGTCCCTTGAAAAAGATTTTCAATTTATTCCTACATTTCTTTCAGCTTATGT 721
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 844 ttggcaaaaagaattctctgtatcagaccagattctcagacaactgttattacacctt 903
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 TTGGCAAGAAATATGTTCTGCCCCACACCTACTTTGATGACTTCTTGGTACCGAAGTGT 781
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 904 gtggccagggtgattctgtatcagattttagtaataatcatcttctctgtggtgattca 963
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 GCTCAGCGGAGGTCTAGATCTCTCTGCGACACACTTTTATTCATCTCTGTGGATTG 841
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 964 acaccaacaatatgaacatgagccgagcaagtgtatatgtgccacacactcttgcggaa 1023
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 ACAAGAAAAAATTAAATGTGAGTCGTTTGTATGTGTACTAGGGCATTAATCCAGCAGGA 901
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1024 catctgtcaaaatattctcaactgagccagcagcagtgaaattctgtgtaactccgggcat 1083
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 CATCTGTTCAAGACTTCTCCACTGGGCACACGCTTGTGTAGATCTGGGAAATTTCAAGCCT 961
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


QY	1142	gtacagagtcagagatagtacggtgccctcacagcaatgtgacagagaggttcagactggct	1201
Db	993	GTACAACATAAAAAACATCGCTTACCCACATGCTTTGTGGAGCGGGCGGGACTGGCT	1052
QY	1202	tccaatccacagagacggtgaataatgctctctcgaagtgaccaaacctcatctaccataa	1261
Db	1053	GGCAGACATCATGACATCATCTTTACTGACTCAATCCCCAAGTTGGTGATCACAA	1112
QY	1262	gaattctcctggaatgggctcacgtggatttcattcctgggggttggatgctctctcacggtat	1321
Db	1113	GAACATTCCTGAGTGGGACCATCTGGATTATTATCTGGGGTTGGATGCCCCCTGGAACT	1172
QY	1322	gtacaatgaatactcatctctgatgcagcagcaggaggag	1358
Db	1173	GTATGATGAATAATTAATCAGTCTAATGAAGAAATACCAG	1209

Search completed: August 24, 2001, 22:44:11
Job time: 9133 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 20:38:04 ; Search time 97.86 Seconds
(without alignments)
3954.138 Million cell updates/sec

Title: US-09-333-159-45

Perfect score: 2044

Sequence: 1 gtcgaccacgctccggg.....aaaaaaaaaaggcgccgc 2044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.2	21.7	1137	1 US-08-227-108-2	Sequence 2, Appli
2	444.2	21.7	1137	2 US-09-073-674-2	Sequence 2, Appli
3	444.2	21.7	1140	1 US-08-227-108-4	Sequence 4, Appli
4	444.2	21.7	1140	2 US-09-073-674-4	Sequence 4, Appli
5	444.2	21.7	1146	1 US-08-227-108-6	Sequence 6, Appli
6	444.2	21.7	1146	2 US-09-073-674-6	Sequence 6, Appli
7	444.2	21.7	1528	1 US-08-227-108-1	Sequence 1, Appli
8	444.2	21.7	1528	2 US-09-073-674-1	Sequence 1, Appli
9	48.2	2.4	178	1 US-08-751-782-2	Sequence 2, Appli
10	48.2	2.4	178	2 US-08-925-171-2	Sequence 2, Appli
11	46.4	2.3	43676	3 US-09-356-952-12	Sequence 12, Appli
12	42	2.1	8041	1 US-08-765-081-1	Sequence 1, Appli
13	42	2.1	8041	3 US-09-098-082-1	Sequence 1, Appli
14	42	2.1	8041	5 PCT-US95-06994-1	Sequence 1, Appli
15	40.4	2.0	1140	2 US-08-698-805-5	Sequence 5, Appli
16	38.6	1.9	2815	1 US-08-230-491A-1	Sequence 1, Appli
17	38.6	1.9	2815	2 US-08-619-280A-1	Sequence 1, Appli
18	38.6	1.9	2815	1 US-08-940-391-1	Sequence 1, Appli
19	38.4	1.9	2798	4 US-09-318-448-34	Sequence 34, Appli
20	36.8	1.8	1248	2 US-08-897-340-5	Sequence 5, Appli
21	36.8	1.8	1248	3 US-09-252-329-5	Sequence 5, Appli
22	36.8	1.8	1406	3 US-08-913-842-6	Sequence 6, Appli
23	36.6	1.8	3271	2 US-08-852-806-1	Sequence 1, Appli
24	36.6	1.8	3271	3 US-08-163-669-1	Sequence 154, Appli
25	36.6	1.8	3282	1 US-08-276-852-154	Sequence 169, Appli
26	36.6	1.8	3282	1 US-08-276-852-154	Sequence 154, Appli
27	36.6	1.8	3282	1 US-08-899-575-154	Sequence 154, Appli

C 28	36.6	1.8	3282	1 US-08-899-575-154	Sequence 169, App
29	36.6	1.8	3282	1 US-08-899-575-154	Sequence 154, App
C 30	36.6	1.8	3282	1 US-08-899-575-154	Sequence 154, App
31	36.6	1.8	3282	5 PCT-US95-08743-154	Sequence 154, App
C 32	36.6	1.8	3282	5 PCT-US95-08743-154	Sequence 154, App
33	36.6	1.8	13254	1 US-08-276-852-156	Sequence 156, App
C 34	36.6	1.8	13254	1 US-08-276-852-156	Sequence 156, App
35	36.6	1.8	13254	1 US-08-899-575-156	Sequence 170, App
C 36	36.6	1.8	13254	1 US-08-899-575-156	Sequence 156, App
37	36.6	1.8	13254	1 US-08-899-575-156	Sequence 170, App
C 38	36.6	1.8	13254	1 US-08-899-575-156	Sequence 156, App
39	36.6	1.8	13254	5 PCT-US95-08743-156	Sequence 170, App
C 40	36.6	1.8	13254	5 PCT-US95-08743-170	Sequence 156, App
41	36.4	1.8	234	3 US-08-651-472-89	Sequence 89, Appli
C 42	35.8	1.8	6027	2 US-08-968-542C-1	Sequence 1, Appli
43	35.2	1.7	37895	1 US-08-375-709-1	Sequence 1, Appli
C 44	35.2	1.7	37895	1 US-08-752-929-1	Sequence 1, Appli
45	35.2	1.7	37895	3 US-09-090-793-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-227-108-2
; Sequence 2, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 21.7%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 1.1e-117;
Matches: 689; Conservative 0; Mismatches 408; Indels 0; Caps 0;

Db	266	ACGCCGGGTACGACGTGTGCTGGGNAACACAGGGGCAACACCTTGGCCACAGGAGATC	325
Qy	550	agacactctccatagaccaagaatgagttctctgggtcttcagttatgatgatgagttgaggt	609
Db	326	TGTACTACTCGCCCGGACTCCGTCGAATTTCTGGCTTTTCAGCTTTTGACGAGATGGCTAAAT	385
Qy	610	ttgacctctctcgagtgaataacttatttgcgagaaacggcgccaggaagaaagatctatt	669
Db	386	ATGACCTTCCGCCACCATTGACTTCATCTTGAAGAAACGGGACAGGCAAGCTACACT	445
Qy	670	atgtcggtatttcacagggcaccaccatttgggtcttatttgcattttccaccatgcacagc	729
Db	446	ACGTTGGCCATTCCCGAGGACACACCATTTGTTTCATCGCTTTTCCACCAATCCCAAGC	505
Qy	730	tggtcagaaaatcaaatgtatttggcttagcaccatagccacitgtaagcatgcaaa	789
Db	506	TGGGAAACGGATCAAAACCTTCATATGATTTAGTCCCGTTGCCACCGTGAAGTACACCG	565
Qy	790	aaagcccccgggaccacaatttttggctgcccagatgatgatcaaggagattgtttggca	849
Db	566	AAACCGTGTAACAAACTCATGCTCGTCCCTTGGTCTCTCTCAAGCTTATATTGGAA	625
Qy	850	aaaaagaattctgtatcagaccagattctcagaaactgttatttaccctttgtggcc	909
Db	626	ACAAATATTCACCCACACCACTCTTTTGATCAATTTCTCGCCACCGAGGTATGCTCCC	685
Qy	910	agtgattctgtatcagattttgtagtaatacatgttacttctctgggtggattcaacacca	969
Db	686	GCAGACGGTGGATCTCCTCTCGACAAACGCCCTGTTTATCATTTGTGGATTTCACACTA	745
Qy	970	acaatatgacatgacgcagcaagtgatatgtgtcccaactcttctgtcgggaacatctg	1029
Db	746	TGAACTTGAAACATGAGTCGTTGGATGTGTATCTGTACATAATATCCACGACGAAACATCGG	805
Qy	1030	tgcataaatattctacactggagccaggcagtgaaattctgttgaactccggcatttgact	1089
Db	806	TTCAGAAACGTGCTCCACTGGTCCCAGGCTGTTAAGTCTGGGAAGTTCCAAGCTTTTGACT	865
Qy	1090	ggggagtgagaccaaaatctgaaaaatcgaatcgaacactcctgtgaaggtacagag	1149
Db	866	GGGAAAGCCCACTTCAGAAACATGATGCATCTATCATATCATAGAGCATGCCCTCTACTACAAC	925
Qy	1150	tcagagatatgacggtccctacagcaatgtggacagagggtccaggactggccttcaaatc	1209
Db	926	TGACAGACATGCATGTGCCAATCGCAGGTGGAACGGTGGCAACGACTTCTTGCCCGACC	985
Qy	1210	cagaagacgtgaaaatgctgtctctctgaggtgacaaacctcatctaccaagaatttc	1269
Db	986	CTCAGATGTTGACCTTTTGCTTTCCAAAGCTCCCAATCTCATTTACGACGGAAGATTC	1045
Qy	1270	ctgaattgggctcacgtgatttcatcttgggtttggatgctccctcaccgatgtacaatg	1329
Db	1046	CTCCTTACAATCATCTGCACCTTATCTCGGGCCATGGATGCCCTTCAAGCGGTTTACAATG	1105
Qy	1330	aaatcatcatctgatg	1346
Db	1106	AAATTGTTTCCATGAT	1122

RESULT 3

RESOLUTION 3
US-08-227-108-4

03-08-227-108-4
; Sequence 4, Application US/08227108

; Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard

APPLICANT: Benicou

APPLICANT: Junien

;	TITLE OF INVENTION:
	Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

QY 790 aaagcccgagcaaaatttttggctgcccagatgatgatcaaggaggtttgttggca 849
Db 569 AAACCCCTGTTAAACAACTCATCTCTGCTCCCTCTGCTCTCTCAAGCTTATATTTGAA 628
QY 850 aaaaagaatttctgtatcagaccagatttccagacaattgttattacatttggcc 909
Db 629 ACAAAATATTCTACCAACACACATCTTTGTATCAATTTCTGCCACCGAGGTATGCTCCC 688
QY 910 aggtgattctgtacagattgttagtaataatctacttctggtgggttcaacacca 969
Db 689 GCGAGACGGTGAATCTCTCTGACAGACGCTCTGTTATCATTTGTGGATTTGACACTA 748
QY 970 acaatatgaacatgagccgagcaagtgtatatgtgtccacacacttctgtggaacattctg 1029
Db 749 TCAACTTGAACATGAGTGGCTGGATGTGTATCTGTACATAATCCAGCAGGAACATCGG 808
QY 1030 tgaataattctacactgagccagcagcagtggaatttctgtggaactccgggcatattgact 1089
Db 809 TTCAGACGCTCTCCACTGGTCCCGAGGCTGTTAAGTCTGGGAAGTTCAGGCTTTTGACT 868
QY 1090 gggggagtgagacaaatctggaataatgcaatcagcccaactcctgttaagggtacagag 1149
Db 869 GGGGAAGCCAGTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACACACC 928
QY 1150 tcaagatatgacggtccctacagaaatgtggacagaggtcagagactggtttcaaatc 1209
Db 929 TGACAGACATGATGTGCAATCGCAGTGTGGAACGGTGGCAAGCACTTGTGGCCGACC 988
QY 1210 cagaagcgtgaaatctgctctgagtgagtgacaaacctcatcaccataagaatttc 1269
Db 989 CTCAGATGTTGACCTTTTGCTTCCAGAGTCCCAATCTCATTTACCCACAGGAAGATTC 1048
QY 1270 ctgaatgggctcagtgattctcctgggtttggatgctcctcaccgtatgtacaaatg 1329
Db 1049 CTCCTTAAACACTTGGACTTTATCTGGCCATGGATGCCCTCAAGCGTTTACAAATG 1108
QY 1330 aaateatccattctgatg 1346
Db 1109 AAATGTTTCCATGATG 1125

RESULT 4

US-09-073-674-4
; Sequence 4, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073.674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-09-073-674-4

Query Match 21.7%; Score 444.2; DB 2; Length 1140;
Best Local Similarity 62.8%; Pred. No. 1.1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 acccagaagcattcatgaatattagtgaaatccatcccaacatcaaggctcatccctgtgagg 309
Db 29 ACCCTGAAGTCACCATGAATATAAGTCAGATGATCACTACTGGGATACCCAGCTGAGG 88
QY 310 aatatgaagtcgaactgaagatgggtatatcccttctgttaacaggagattcctcagagcc 369
Db 89 AATATGAAGTGTGACCGAAGACGTTATATCTTGGATCGACAGAAATCTCTTATGGGA 148
QY 370 tagtgcaactaagaagacagaggttccagggcctgtgggttactcagcagatggcctagtgtg 429
Db 149 GGAAAAATTCAGAGAAATATAGGCGCGAGACCTGTTCGATTTTTCACACAGGTTTGCTCG 208
QY 430 gagtggctagcaactgattccaaacctgcccacaaatagcctgggcttctcattctcggcag 489
Db 209 CATCAGGCCACAACTGATCTCCAACTGCGCCCAACACAGCGCTGGCCCTTCATCTCTGGCCG 268
QY 490 atgtgtgttttgacgtgtggatggggaaacagcaggggaaacagcctgggtctcgaaaaacaca 549
Db 269 ACGCGGGGTACGACGCTGTGGCTGGGAACAGCAGCGGGCAACACCTGGCGCAGGAGGAATC 328
QY 550 agacactctccatagacacagatgagttctgggttttaagtattatgatgagatggctaggt 609
Db 329 TGTACTACTCGCCGCACTCCGTCGAATTTCTGGGCTTTTCAGCTTTTACAGAGATGGCTAAAT 388
QY 610 ttgacctctcgtcagtgataaaactttatttgcagaaaacggccagagaaagatctatt 669
Db 389 ATGACCTTCCCGCCACCATTCATCTTTGAAGAAAACGGGACAGACAGCTACACT 448
QY 670 atgtcggctattcagagggcaccaccattgggctttattgcattttccaccatgcccagagc 729
Db 449 ACGTGGCCATTCCAGGGCACCACTTGGTTTTCATCGCTTTTCCACCAATCCCAAGC 508
QY 730 tggctcagaaaaatcaaaatgatttctgttagcaccatagccactgttaagcatgcaa 789
Db 509 TGGGAAACGGGATCAAAACCTTCTATGATTTAGTCTCGGTTGCCACCTGAGTACACCG 568
QY 790 aaacccccgggaccaaattttgtgtgcccagatgatgatcaaggagattgtttggca 849
Db 569 AAACCTGTTAAACAACTCATGCTCGTCCCTCTGTTCTCTTCAAGCTTATATTGGA 628
QY 850 aaaaagaatttctgtatcagaccagatttctcagacaacttgttattacatttggcc 909
Db 629 ACAAAATATTCTACCAACACACCTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 688
QY 910 aggtgattctgtacagattgttagtaataatcatttacttctgtggtgagttcaacacca 969
Db 689 GCGAGCGGTGGATCTCTCTGACAGCAACGCCCTGTTTATCATTTGTGGATTTGACACTA 748
QY 970 acaatatgaacatgagccgagcaagtgtatatgtgtccacacacttctgtggaacattctg 1029
Db 749 TGAACCTTGAACATGAGTGGCTTGGATGTGTATCTGTCACATANTCCACAGGAGACATCGG 808
QY 1030 tgcataattctacactgagccagcagtgaaattctgttgaactcctggtgagcgggacttgact 1089

Db 809 TTCAGACGCTGCTCCACTGGTCCAGGCTGTAAGCTCGGAGTTCACAGCTTTTGACT 868
Qy 1090 gggggagtgagaccacaaatctggaaaaatgcaatcagccaaactcctgtaaggtagag 1149
Db 869 GGGGAAGCCAGCTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAACC 928
Qy 1150 tcagagatatgagctccctcagcaaatgtggacagagggtcagagactggctttcaaatc 1209
Db 929 TGACAGACATGATGTCATCGCAATCGCAGTGTGGAACGGTGGCAACACATGCTGGCCGACC 988
Qy 1210 cagaagacgtgaaatgctgctctctgagtgagacacaaactcatctaccataagaatattc 1269
Db 989 CTCAGATGTTGACCTTTTGCTTTCCAGCTCCCAATCTCATTTACACAGGAGATTC 1048
Qy 1270 ctgaatgggctcactgagatttcctctgggttggatgctcctcactgagatg 1329
Db 1049 CTCCTTACAATCACTTGGACTTTATCTGGGCATGGATGCCCTCAAGCGGTTTACAATG 1108
Qy 1330 aaatcatccatctgatg 1346
Db 1109 AAATGTTTCCATGATG 1125

RESULT 5

US-08-227-108-6

; Sequence 6, Application US/08227108

; Patent No. 5807726

; GENERAL INFORMATION:

; APPLICANT: Blanchard, Claire

; APPLICANT: Benicourt, Claude

; APPLICANT: Junien, Jean-Louis

; TITLE OF INVENTION: Recombinant Dog Gastric Lipase

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/227,108

; FILING DATE: 03-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fanucci, Allan A.

; REGISTRATION NUMBER: 30,256

; REFERENCE/DOCKET NUMBER: 7620-033

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212 790-9090

; TELEFAX: 212 869-8864/9741

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1146 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

US-08-227-108-6

Query Match

Best Local Similarity 21.7%; Score 444.2; DB 1; Length 1146;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 250 acccagaagcattcatgaattagtgaaatcaccacatcaaggctatccctgtgagg 309

Db 26 ACCCTGGAAGTGACCATGAATATAAGTCAGATGATCACCTACTGGGGATACCCAGGTGAGG 85
Qy 310 aatagaagtgcgaactgaagatgggtatatactcttcttgtaacaggaattctctcagggcc 369
Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTCTTGGGATCCACAGAAATTCCTTATGGGA 145
Qy 370 tagtgaacacctaaagaacagagttccagggcctgtgttactcagcaatggtccttagttg 429
Db 146 GGAAAAATTCAGAGAAATATAGCCCGGAGACCTGTTCATTTTGAACAACGGTTTGCTCG 205
Qy 430 gaggtgctagcaactgagatttccaaacctgcccacaaatagcctgggtcttcattctggcag 489
Db 206 CATCAGCCACAAACTGGGATCTCCAACCTGCCCAACAACAGCCTGGCCCTCATCTCTGGCCG 265
Qy 490 atgctgggtttgacgtgtggatggggagacagcaggggaaacacccctggtctcgaacacaca 549
Db 266 ACGCCGGGTACGACGTGTGGCTGGGGAACAGCAGGGGCAACACCTGGGCCAGGAGAAATC 325
Qy 550 agacactctccatagacccaagatgagttctgtggttccagtttatgatgagatggttaggt 609
Db 326 TGTACTACTCGCCGACTCCCGTCGAATTTCTGGCTTTTTCAGCTTTGACGAGATGGCTAAAT 385
Qy 610 ttgacctctcctgcagtgataaaactttatttgcagaaaaacggggccaggaagaaatctatt 669
Db 386 ATGACCTTCCCGCCACCATTGACTTCATCTTGAAGAAAACGGGACAGGAACTACACT 445
Qy 670 atgtcggtattcacagggccaccacacacacacacacacacacacacacacacacacacac 729
Db 446 ACGTTGGCCATTCACGAGGACCACTGTTTTCATCGCTTTTCCACCAATCCCAACG 505
Qy 730 tggctcagaaaaatcaaaatgtattttgttttagccacacacacacacacacacacacacac 789
Db 506 TGGGAAACGGATCAAAACCTTCTATGATTAAGCTCCCGTTGCCACCGTGAAGTACACCG 565
Qy 750 aagccccgggacacaaattttgtgtgcagacacacacacacacacacacacacacacacac 849
Db 566 AAACCTTAAACAACTCATGCTCGCTCCCTTCTCTTCAAGCTTATATTGGAA 625
Qy 850 aaaaagaatttctgtatcagaccagatcttctcagacacacacacacacacacacacacac 909
Db 626 ACAAAATATTCTACCCACACACACTTCTTTGATCAATTTCTCCGCCACCGAGGTATGCTCC 685
Qy 910 agg-gattctctgacagattgttagtaataatcttactctctggtggtggttaacacaca 969
Db 686 GCGAGACGTTGGATCTCTCTGACGACACGCGCTTTTATCATTTTGTGGATTGACACTA 745
Qy 970 acatataaacatgagccgagcaagtgtatatgtgcacacacacacacacacacacacacac 1029
Db 746 TGAACCTTGAACATGAGTCGCTTGGATGTGTATCTCTGTCAATATCCAGCAGGAACATCG 805
Qy 1030 tgcnaaatattctacactggagccagggcagtgaaatctcgtgtaactccgggctatttgact 1089
Db 806 TTCAGAACGCTGCTCCACCTGGTCCCGAGGTGTTAAGTCTGGGAAGTTCGAAGCTTTTGACT 865
Qy 1090 gggggagtgagacccaataatctggaaaaatgcaatcagccaaactcctgtaaggtagag 1149
Db 866 GGGGAAGCCAGTTTCAGAACATGATGCACTATCATCAGAGCATGCTCTCCCTACTACAAC 925
Qy 1150 tcagagatatgagcgtccctcagcaaatgtggacaggggtcagagactggcttcaaatc 1209
Db 926 TGACAGACATGATGTCGAATTCGCAATTCGGAACGGTGGCAACGACTTGTCTGGCCGACC 985
Qy 1210 cagaagacgtgaaatgctcctctgagtgagacacaaactcatctaccataagaatattc 1269
Db 986 CTCAGATGTTGACCTTTTGCTTTTCCAAAGCTCCCCAAATCTATTACCACAGGAGATTC 1045
Qy 1270 ctgaatgggtcagtgagatttcattctctgggttttggatgctcctcaccgtatgtacaa 1329
Db 1046 CTCCTTACAATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAAGCGGTTTACAATG 1105
Qy 1330 aaatcatccatctgatg 1346

Db 1106 AATTGTTCCATGATG 1122

RESULT 6
US-09-073-674-6
; Sequence 6, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-09-073-674-6

Query Match 21.7%; Score 444.2; DB 2; Length 1146;
Best Local Similarity 62.8%; Pred. No. 1.1e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 accagaagcattcatgaattagtgaaatcatcacaatcaagaagcctccctgtgag 309
|||||
Db 26 ACCCTGAAGTGACCAATGAATAGTCAGATGATACCTACTGGGATACCCAGCTGAGG 85
|||||

QY 310 aatgaagtcgcaactgaagatggtatctcttcttcttcaacagattccctgagcc 369
|||||
Db 86 AATATGAAGTTGTGACCGAAGCGGTATATCTCTGGGATCGACAGAAATCTCTATGGGA 145
|||||

QY 370 tagtgaacctgaagaagacaggtttccaggcctgtgtgttactgcagcagatgacctagt 429
Db 146 GGAATAATTCAGAGATATAGCCGCGAGACCTGTTGCATTTTGCACACGGTTTCTCTG 205
|||||

QY 430 gaggtgctagcaactgagatttccaaactgcccacaatagcctgggcttcttctgagcag 489
|||||

Db 206 CATCAGCCACAACTGGATCTCCAACTGCCAACAAACAGCTGGCTTCATCTCTGCGCG 265
|||||

QY 490 atgctggtttgagtggtgaggggaacagcaggggaacgctggtctcgaacacaca 549
|||||

Db 266 ACCCGCGGTACGACGTGTGGCTGGGAACAGCAGGGGGCAACACCTGGCGCAGGAGGAATC 325
|||||

QY 550 agacactcttcacagacgaagatggttctggtcttcagttatgatgatgagtggtcag 609
|||||

Db 326 TGTACTACTGCGCCGAGCTCGTGGAAATTCCTGGGCTTTTCAGCTTTGACGAGATGCTAAAT 385
QY 610 tgaaccttctcagtgagataaaactttattttgcagaaaaacggccaggaagaaagatctatt 669
|||||
Db 386 ATGACCTTCCCGCCACCATTTGACTTTTGAAGAAAACGGGACAGGAAAGTACACT 445
QY 670 atgctggtctattcacaggggaccacacatgggctttattgcatttccaccattgcccagc 729
|||||
Db 446 ACCTTGGCCATTCACAGGACACCATTTGTTTCATCGCTTTTCCACCAATCCCAGC 505
QY 730 tggctcagaaaaataaatagtatttggcttttagcaccatagccactgttaagcatgcaa 789
|||||
Db 506 TGGCGAAACGGATCAAAACCTTCTATGCATTAGCTCCGTTGCCACGCTGAAGTACACCG 565
QY 790 aaagcccccgggaccaaaatttttctgctccagatatgatgatacaaggattgtttggca 849
|||||
Db 566 AAACCTGTTAACAACACTCATGCTGCTCCCTTCCTGCTTCCTTCACCTTATATATGGAA 625
QY 850 aaaaagaatttctgtatcagaccagatttctcagacaaactgttattaccctttgtggcc 909
|||||
Db 626 ACAATAATTTCTACCCACACCATTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 685
QY 910 aggtgattcttgatcagattgttagtaataatcatgttactctctggtggattcaacacca 969
|||||
Db 686 GCGAGACGGTGGATCTCTCTGCGCAACACCCCTGTTTATCATTTGTGGATTGTGACACTA 745
QY 970 acaatgaacatgagccgagcaagtgatatgtcgtccacacactctgtggaacacatctg 1029
|||||
Db 746 TGAACCTTGACATGAGTCGCTTGGATGTGTATCTGTGCATATATCCAGCAGGAACATCGG 805
QY 1030 tgcataaattctacactggagccagcagtgatcttctgtgaactccggcatttgact 1089
|||||
Db 806 TTCAAGAACGTGCTCCACTGGTCCAGGCTGTTAAGTCTGGGAAGTTCNAAGCTTTTGACT 865
QY 1090 gggggagtgagacacaaaactctggaaaaatgcaaatcagccaaactctgttaaggtacagag 1149
|||||
Db 866 GGGGAAGCCCGAGTTCAGAACATGATGCACTATCATCAGACATGCTCCCTACTACAAAC 925
QY 1150 tcagagatatgacggtccctacagcaatgtggacagagagtcagagctggtcttcaaatc 1209
|||||
Db 926 TGACACATGATGATGTCCTGTCGAGTGTGGAAACGGTGGCAACGACTTCTGCGCCGACC 985
QY 1210 cagangacgtgaaaaatgctgctctctgaggtgacacacactcatctaccataagaatctc 1269
|||||
Db 986 CTATCATGTTGACCTTTTGTCTTCAAGCTCCCAATCTCATTTACACAGGAAGATTC 1045
QY 1270 ctgaatgggctcagtggtatttcatctgggttttggatgctctcaccgtatgtacaatg 1329
|||||
Db 1046 CTCCTTACAACTACTTGGACTTTTATCTGGGCGCATGGATGCCCTCAAGCGGTTTACAATG 1105
QY 1330 aaatcatccatctgtatg 1346
|||||
Db 1106 AATGCTTTCCATGATG 1122

RESULT 7
US-08-227-108-1
; Sequence 1, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:

QY	1210	cagaagacgtagaaatctgctctctctgaggtgagcacaacctcatctaccataagaatattc	1269
Db	986	CTCAGATGTGACCTTTTGGCTTTCCAAAGCTCCCAATCTCATTATCCACAGGAAGATTC	1045
QY	1270	ctgaatgggctcagctggtattcatctggtggttggatgctcctcacccgtatgtacaaatg	1329
Db	1046	CTCCTTACATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAACGGCTTACAAATG	1105
QY	1330	aaatcatccatctgtatg	1346
Db	1106	AAATGTTCATGATG	1122
RESULT 9			
US-08-751-782-2			
; Sequence 2, Application US/08751782			
; Patent No. 5821352			
GENERAL INFORMATION:			
; APPLICANT: Reintz, Nathaniel			
; APPLICANT: Gubbay, Johnathan			
; APPLICANT: Skinner, Michael			
; TITLE OF INVENTION: A CDNA Library Prepared during			
; NUMBER OF INVENTIONS: Regression of Rat Prostate and Methods of Use Thereof			
; NUMBER OF SEQUENCES: 15			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: David A. Jackson, Esq.			
; STREET: 411 Hackensack Ave, Continental Plaza, 4th			
; STREET: Floor			
; CITY: Hackensack			
; STATE: New Jersey			
; COUNTRY: USA			
; ZIP: 07601			
COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/751,782			
; FILING DATE: 18-NOV-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Jackson Esq., David A.			
; REGISTRATION NUMBER: 26,742			
; REFERENCE/DOCKET NUMBER: 600-1-190			
TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 201-487-5800			
; TELEFAX: 201-343-1684			
; INFORMATION FOR SEQ ID NO: 2:			
SEQUENCE CHARACTERISTICS:			
; LENGTH: 178 base pairs			
; TYPE: nucleic acid			
STRANDEDNESS:			
; TOPOLOGY: double			
; MOLECULE TYPE: cDNA			
; HYPOTHETICAL: NO			
; IMMEDIATE SOURCE:			
; CLONE: 10.2			
US-08-751-782-2			
Query Match 2.4%; Score 48.2; DB 1; Length 178;			
Best Local Similarity 55.8%; Pred. No. 0.00013;			
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0			
QY	541	gaaacacaaagacacactctccatagaccagaatgagttctgggctttcagttatgatgaga	600
Db	8	GGAAACATGTGGCCTACACCCAGATTTCTAAAGATTTTGGGATTTTAGTTTTAATGAAC	67
QY	601	tggcagggttgaccttctcctgcagtgataaacctttatttgcagaaaaacggccaggaaa	660
Db	68	AAATAGAATAGACACCTCCGACGCATCATTTATTTCATTCGTAATGAACAGACAACAC	127

ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44Mb storage
COMPUTER: IBM PC compatible/Pentium
OPERATING SYSTEM: MS-Windows 3.1
SOFTWARE: Word for Windows-6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,081
FILING DATE: March 26, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06994
FILING DATE: June 7, 1995
APPLICATION NUMBER: US 08/265,714
FILING DATE: June 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K.
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: CHOR-1-10286
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-682-8100; 1-206-224-0718 (direct)
TELEFAX: 1-206-224-0779
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli O157:H7
STRAIN: 86-24 NALR
CLONE: PEAR
IMMEDIATE SOURCE:
US-08-765-081-1

Query Match 2.1%; Score 42; DB 1; Length 8041;
Best Local Similarity 53.4%; Pred. No. 0.058;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 33 aaaaatgtgaagagtttttaaaacccacacaaattcttcttacttttagaattagttgttaca 92
Db 1541 AAAGTATGTGAAGCTTTATTTTAAACACACAGATAATTGATTTTAAAAATAAATGTTAGA 1482
QY 93 ttggcagggaagaaataaaatgcagatgttgaccatgttgaaacctgtgcaagacagtg 152
Db 1481 TTCCAGAAATATTTAAGAATAACCGACATTAACAGGTAAGCTGTATGGCTATGG 1422
QY 153 atgtctcacagaaatggaatgtgcttctgattctgtg 195
Db 1421 ATATGNCATTTGTTTAAATGTTAAATGTTGTTGTTGG 1379

RESULT 13
US-09-098-082-1/c
Sequence 1, Application US/09098082
Patent No. 6040421
GENERAL INFORMATION:
APPLICANT: Tarr, P.I., Bilge, S.S., Besser, T.E., Vary Jr., J.C.
TITLE OF INVENTION: Escherichia Coli O157:H7 Epithelial Adhesin
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Center, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44Mb storage
COMPUTER: IBM PC compatible/Pentium II
OPERATING SYSTEM: MS-Windows 95
SOFTWARE: Word for Windows-6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,082
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/765,081
FILING DATE: March 26, 1997
APPLICATION NUMBER: PCT/US95/06994
FILING DATE: June 7, 1995
APPLICATION NUMBER: US 08/265,714
FILING DATE: June 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Shelness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: CHOR-1-12402
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-682-8100; 1-206-224-0735 (direct)
TELEFAX: 1-206-224-0779
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli O157:H7
STRAIN: 86-24 NALR
CLONE: PEAR
IMMEDIATE SOURCE:
US-09-098-082-1

Query Match 2.1%; Score 42; DB 3; Length 8041;
Best Local Similarity 53.4%; Pred. No. 0.058;
Matches 87; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 33 aaaaatgtgaagagtttttaaaacccacacaaattcttcttacttttagaattagttgttaca 92
Db 1541 AAAGTATGTGAAGCTTTATTTTAAACACACAGATAATTGATTTTAAAAATAAATGTTAGA 1482
QY 93 ttggcagggaagaaataaaatgcagatgttgaccatgttgaaacctgtgcaagacagtg 152
Db 1481 TTCCAGAAATATTTAAGAATAACCGACATTAACAGGTAAGCTGTATGGCTATGG 1422
QY 153 atgtctcacagaaatggaatgtgcttctgattctgtg 195
Db 1421 ATATGNCATTTGTTTAAATGTTAAATGTTGTTGTTGG 1379

RESULT 14
PCT-US95-06994-1/c
Sequence 1, Application PC/TUS9506994
GENERAL INFORMATION:
APPLICANT: Children's Hospital & Medical Center
APPLICANT: University of Washington
APPLICANT: Washington State University Research Foundation
APPLICANT: TARR, PHILLIP I
APPLICANT: BILGE, SIMA S
APPLICANT: BESSER, THOMAS E
APPLICANT: VARY JR, JAMES C
TITLE OF INVENTION: ESCHERICHIA COLI O157:H7 EPITHELIAL ADHESIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON AND KINDNESS/PLLC

STREET: SUITE 2800, 1420 FIFTH AVENUE
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: WA 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Vers
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06994
FILING DATE: 07-JUN-95
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/265,714
FILING DATE: 24-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: BRODERICK, THOMAS F
REGISTRATION NUMBER: 31,332
REFERENCE/DOCKET NUMBER: CHOR-18591
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682 8100
TELEFAX: (206) 224 0779
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Escherichia coli Q157:H7
STRAIN: 86-24 NALR
IMMEDIATE SOURCE:
CLONE: pEAR
PCT-US95-06994-1

	Query Match	2.1%	Score 42;	DB 5;	Length 8041;
	Best Local Similarity	53.4%;	Prod. No. 0.058;		
	Matches 87;	Conservative 0;	Mismatches 76;	Indels 0;	Gaps 0;
QY	33	aaaatatgtgaagagctttttaaaccacaaattctcttactacttagaataattgtgtaca	92		
Db	1541	AAAGTATGTGAACCTATTATTTTAAACACACAGATAATTGTATTTTAAAAATAAATGTATGA	1482		
QY	93	tgggcaggagaaaaataaatgcagatgttggaccactgttggaaacctgttcaagacacagtg	152		
Db	1481	TTTCCAGAAATAATTTTAAAGAAATACCGACATTACAAAGGTGAAGTCGTATGGCTATGG	1422		
QY	153	attgtctcacagagaatggaaatgtggctcttgattctctgttg	195		
Db	1421	ATATGNCATTTGTTATTGAATGGTGAATTTGTTGTTACTTTGG	1379		

RESULT 15
US-08-698-805-5
: Sequence 5, Application US/08698805
: Patent No. 5869288
: GENERAL INFORMATION:
: APPLICANT: Chapman, Martin
: APPLICANT: Arruda, L. Karla
: TITLE OF INVENTION: Molecular Cloning of Cockroach
: TITLE OF INVENTION: Allergens, Amino Acid and Nucleotide Sequences Therefore.
: TITLE OF INVENTION: and Recombinant Expression Thereof
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Oblon, Spivak, McCrelland, Maier & Neustadt,
: ADDRESSEE: P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/698,805
FILING DATE: 16-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,510
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kelber, Steven B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 494-203-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..602
PS-08-698-805-5

Query Match	2.08;	Score 40.4;	DB.2;	Length 1140;
Best Local Similarity	45.5%;	Pred. No.	0.059;	
Matches 143;	Conservative	0;	Mismatches 171;	Indels 0; Gaps 0;

Qy	1525	gaaccttacattacatcttttctgtaaatgaagcacttataggtaataagaggtt	1584
Db	827	GACGCTGATTCCTACATTAAATCTTCACAATTTAGAAGTTTTTAACAATAGTAATTAC	886
Qy	1585	ttg:atgcctattatatctaccatctctgaagggtaggttttacctgatgcagcaaaa	1644
Db	887	GATUACAATTTACAGATCTGTTAGATATGATTCGAAGCTGTGTTATAANTCAGAAAAATGA	946
Qy	1645	tat:tagacatctcttatatacttcaggtaaaatctctttaaacacccatctgtttttct	1704
Db	947	CTTGCTAAAATCAAACAGCATATGCGCAAAATTTTTTCGTTCTAANNTCCAGTTTTTTAA	1006
Qy	1705	ataagcccatattttgagcacctaagaatcgcaaatggagacagatatagggtct	1764
Db	1007	ATG'PATAAATTTTTTGGTAAACTTTATTTACTAGAAATTTGATCCAGAGTAGACTGAT	1066
Qy	1765	ggagtctgtgatttatgttgttcatttgacaaaaataagctagacattttccacctgtgcc	1824
Db	1067	AATITCCTTTACTTACTTTTTTGGTATTAAACAAAGTTGGAAACAAAAATTTTCAAAA	1126
Qy	1825	acagagacataaca	1838
Db	1127	AAAAAAAAAAAAAA	1140

Search completed: August 24, 2001, 22:38:56
Job time: 7252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 19:50:13 ; Search time 2518.61 Seconds
(without alignments)
7671.546 Million cell updates/sec

Title: US-09-333-159-45
Perfect score: 2044
Sequence: 1 gtcgaccacgcgtccggg.....aaaaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	628.2	30.7	2927	192	AK019504	Mus muscu
2	505	24.3	954	144	BF135102	601780142
3	497.6	24.3	1079	145	BF141567	601788654
4	492	24.1	669	145	BF143546	601790896
5	461.8	22.6	1350	192	AK010093	Mus muscu
6	460.2	22.5	1347	192	AK010116	Mus muscu
7	460.2	22.5	1349	192	AK009300	Mus muscu
8	460.2	22.5	1350	192	AK010139	Mus muscu
9	458.6	22.4	1340	192	AK010124	Mus muscu
10	458.6	22.4	1343	192	AK009479	Mus muscu
11	458.6	22.4	1345	192	AK009413	Mus muscu
12	458.6	22.4	1345	192	AK009473	Mus muscu
13	458.6	22.4	1346	192	AK009359	Mus muscu
14	458.6	22.4	1346	192	AK009523	Mus muscu
15	458.6	22.4	1346	192	AK009729	Mus muscu
16	458.6	22.4	1348	192	AK010058	Mus muscu
17	458.6	22.4	1349	192	AK009474	Mus muscu
18	458.6	22.4	1349	192	AK009546	Mus muscu
19	458.6	22.4	1349	192	AK009571	Mus muscu
20	458.6	22.4	1349	192	AK009573	Mus muscu
21	458.6	22.4	1349	192	AK010148	Mus muscu
22	458.6	22.4	1350	192	AK009459	Mus muscu
23	458.6	22.4	1350	192	AK009525	Mus muscu
24	458.6	22.4	1350	192	AK009773	Mus muscu
25	458.6	22.4	1350	192	AK010061	Mus muscu
26	458.6	22.4	1350	192	AK010125	Mus muscu
27	458.6	22.4	1350	192	AK010203	Mus muscu
28	458.6	22.4	1351	192	AK009428	Mus muscu
29	458.6	22.4	1352	192	AK009431	Mus muscu
30	458.6	22.4	1356	192	AK010019	Mus muscu
31	458.6	22.4	1356	192	AK010035	Mus muscu
32	457.6	22.4	1346	192	AK009437	Mus muscu
33	457.6	22.4	1347	192	AK010103	Mus muscu
34	457.6	22.4	1349	192	AK010026	Mus muscu
35	457.6	22.4	1346	192	AK009560	Mus muscu
36	457	22.4	1347	192	AK010236	Mus muscu
37	457	22.4	1348	192	AK010231	Mus muscu
38	457	22.4	1349	192	AK009875	Mus muscu
39	457	22.4	1350	192	AK009537	Mus muscu
40	456	22.3	1349	192	AK010106	Mus muscu
41	455.4	22.3	1345	192	AK009544	Mus muscu
42	455.4	22.3	1349	192	AK009990	Mus muscu
43	453.8	22.2	642	145	BF138035	601784732
44	453.8	22.2	1346	192	AK010173	Mus muscu
45	450.6	22.0	933	145	BF143673	601789739

ALIGNMENTS

RESULT	1	
AK019504		
LOCUS	AK019504	2927 bp
DEFINITION	Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23, full insert sequence.	HTC 08-FEB-2001
ACCESSION	AK019504	
VERSION	AK019504.1	GI:12859754
KEYWORDS	CAP trapper.	
SOURCE	Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:4632427C23.	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
REFERENCE	1 (sites)	
AUTHORS	Carninci, P. and Hayashizaki, Y.	
TITLE	High-efficiency full-length cDNA cloning	

Query Match	30.7%	Score 628.2	DB 192	Length 2927
Best Local Similarity	84.3%	Pred. No. 1.3e-149		
Matches 734	Conservative 0	Mismatches 128	Indels 9	Gaps 2
QY 572	tgagttctggttcagttcatgatgagatggcttagtttgaccttctcagtgataaa	631		
DB 1376	TGTCCTTTCCCTTTGTAGTTATGATGAATGGGTAGTGTTCAGCTTCCAGCTGTGATAA	1435		
QY 632	ctttattttgcagaaaaacggccaggaagaatctattatgctggtctaticacagggcac	691		
DB 1436	CTTTATCTTACAGAAAACGGCCAGAAAAGGCTATTATGTCGGCTACTCACAGGGCAC	1495		
QY 692	caccatggggtttattgcattttccaccatgcagagactggctcagaaaaatcaaaatgta	751		
DB 1496	CACCATGGGCTTTATTTCGATTTTCCACAAATGCCAGAGCTAGCTCATAAAATCAAAAATGA	1555		
QY 752	ttttgctttagcaccatagccactggttaagcatgcgaaaagcccgccagccaaattttt	811		
DB 1556	TTTTGCCCTTAGCTCCTATAGCCACTGTTAAATATGCAAGAGCTCTGGTACCAAAATTTCT	1615		
QY 812	gttgctgcagatgatgatcatcaagggtattgtttggccaaaaaagaatttctgtatcagac	871		
DB 1616	GCTGCTGCCAGATGATGATGATCAAGGTATTATTGGCAGACAAGATTTTATACCCAGAC	1675		
QY 872	cagatttctcagacaactgttatttaccctttgtgcccaggtgattcttcattcagattg	931		
DB 1676	TAGATTTTTCAGACAGCTTTTATTATACCTTTTGGGCCAGATGATTTTGACCAAAATCTG	1735		
QY 932	tagtaatatcatgttactctgtgggtgattcaacacacacaatatagcattgagccgagc	991		
DB 1736	CAGCAACATCATCTTACTCTCGTGGAGGCTTTAAACACAAAACAATATGAACATGAGCAGAGC	1795		
QY 992	aagtgtatatgctgcacacactctgtgaaactctgtgcaaatcttctacactggag	1051		
DB 1796	AAATGTGTATGTGCCCATACACTGCTGGAAGCTGTGCGAGAACATTTCTCACTGGAG	1855		
QY 1052	ccaggcagtgaaattctgttgaaactccgggcatttgactgggggagtgagaccacaaatct	1111		
DB 1856	CCAGGCCGTGAATCTCGGGAACTTCGTGCTTTGACTGGGAGTCAGACCAAAAATCA	1915		
QY 1112	ggaaaaatgcaatcagcccaactcctgtaaaggtcacagagtcagagatagcaggtccctac	1171		
DB 1916	GGAGAAATGTAATCAGGCCAACTCCTATPAGGTACAAAGTTTCGAGATATGATGTTGCCAAC	1975		
QY 1172	agcaatgtgcacagaggttcaggactggtctcaaatccagaaagcgtgaaatgctgct	1231		
DB 1231	AGCAATGTGCACNTGGAGGTCAGACNTGGCTTTCAATCCAGATGATGTGAAAACATTTACT	2035		
QY 1232	ctctgaggtgaccaacctcatctaccataagaatatcttgaatgggctcacgtggattt	1291		
DB 2036	TTCTGAAGTAAACCAACCTCATCTACCAACAAGACATTCTCTGAGTGGGCTCATGTGGATT	2095		
QY 1292	catctgggggttgatgactcctcacqjattgatacatgaataatccatctgatgcagca	1351		
DB 2096	CATCTGGGGGCTGGATGGCCCTCAGCGTGTTTACATGAATATATACATCTGATGAAGC -	2154		

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 669)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
cDNA Library Preparation: Gilbert Smith, Ph.D.
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9277 row: 6 column: 23
High quality sequence stop: 659.
Location/Qualifiers
1..669
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4021870"
/clone_lib="NCI-CCGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dt. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 195 a 140 c 165 g 169 t
ORIGIN

Query Match 24.1%; Score 492; DB 145; Length 669;
Best Local Similarity 85.5%; Pred. No. 5.9e-115;
Matches 571; Conservative 0; Mismatches 95; Indels 2; Gaps 2;

QY 82 tagttgtacattgagcagaa-aaataaatgcagatgttgaccattgtgaaaccttg 140
|||||
Db 3 TAGTGTTCATCTTATGACAAAGCAATGCCGATGTGGACCATGTCGAAATCTTG 62

QY 141 tcaagacagtgattgtctcacacagaatggaaatgtggctctgattctgtggcgctat 200
|||||
Db 63 TCAAGAGTGTGAGCTCTTTCGCACAGAGTGGAGATATGCTCTGATTCGTGACATAT 122

QY 201 atgttccagagaatgtgaattcagatcatatgcacactaaagctgtggacccagaagca 260
|||||
Db 123 TTAACCAAGAAATGTGAACCTGGGACATTTGCCACGAAAGCTGCGGATCCAGAAGCA 182

QY 261 ttcataatattagtgaaatcccaatcaagctatccctgtgaggaatgaagtc 320
|||||
Db 183 TTCATGAATGTAGCCAAATATCAACAAGGGTTATCCCGAGGAGTGAAGTT 242

QY 321 gcaactgaagatgggtatctcttctgttaacaggattccctcgagccctagtcacact 380
|||||
Db 243 GCAACAGAGATGGGTACATCTTCTGTGACAGAGATCCCTCGGGGACACAGGTTA 302

QY 381 aagaagacaggttccaggcctgtgtgttactgcagcatggcctagtgaggtgctagc 440
|||||
Db 303 AAGAAGAGGATCCAGGCGAGTGGTGTACTGCAACATGGTCTCTTGGGGATGCTAGC 362

QY 441 aactgatttccaaacctgcacacaaatagctgggcttctctgagagatgctggtttt 500
|||||
Db 363 AACTGATTTCACCTGCCCCAACACACACCTGGGTTTATTCGTGCGAGATCGAGG-TTT 421

QY 501 gacgtgtgagtgaggacagcaggggaaacgcctgtgtctcgaacacacacacacctcc 560
|||||
Db 422 GATGTGTGATGGGAACAGCAGAGGAAACACCTGGTCTCGGAAGCACAAAGACCTCTCC 481

QY 561 atagaccaagatgagttcttgggcttttcagttatgatgagatggagtggttgcaccttc 620

Db 482 ATAGTCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAATGGTGGTTCACCTTCCA 541
|||||
QY 621 gcagtgataaacttatttgcagaaaacggccaggaagaaagatctattatcgcgtat 680
|||||
Db 542 GCTGTGATAACTTATCTTACAGAAAACGGCCAGAAAAGGTCTATTATCTCGGCTAC 601
|||||
QY 681 tcacugggcac 740
|||||
Db 602 TCACAGGSCAC 661
QY 741 atcanaat 748
|||||
Db 662 ATCCAAAT 669
|||||
RESULT 5
AK010093 1350 bp mRNA HTC 08-FEB-2001
LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DEFINITION library, clone:2310067K20, full insert sequence.
ACCESSION AK010093
VERSION AK010093.1 GI:12845297
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2310067K20.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000).
MEDLINE 20499374
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T.,
Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T.,
Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000).
MEDLINE 20530913
REFERENCE 4 (sites)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1350)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Haragaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,
Imoto,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Saito,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Socabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Taka,T., Teijima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome

QY	850	aaaagaattctgtatcagaccagagatttctcagacaacttgttattactcttggcc	909
Db	732	ACAAATGTTATGCCCCCAACTACTTAGATCAATTTCTTGGTACGGAAGTGTCAC	791
QY	910	agtgatttctgatacagattgttagtaataatctgttactcttggtggattccaccca	969
Db	792	GGGAGCTGTAGATCTTCCTGCGAGCAACGCTTTATTCATCTTCTGTGGATTGACAAGA	851
QY	970	acaatatgaatgagccgagcaagtgatatctgctcccaactcttgctgggaacatctg	1029
Db	852	AAACTTAAATGTGAGTCGCTTTGATGTGTATCTAGGGCATATCCAGCAGGAACATCTA	911
QY	1030	tgcataattctacatcgagccgagcagtgtaattctgtgaactccggtgact	1089
Db	912	CTCAAGACCTTTTCCACTGGGCACAGCTTCTAAATCTGGGAAGCTTCAAGCCTATAACT	971
QY	1090	ggggagtgagaccacaaaatctctggaaaaatgcaatcagaccaactcctgtgaagtcacag	1149
Db	972	GGGAGTCCATTACAGAACATGTTACACTACATCAGAAACGCCCTCCCTACTACTGATG	1031
QY	1150	tcajagatatgacggtccctcacagcaatgtggacagagctcagagctgcttccaatc	1209
Db	1032	TGTACCCATGACCGTGCCCAATTGCTAGTGTGGAAAGGTGGCCATGACATCCTGGCTGATC	1091
QY	1210	cagagacgtgaaatctgctctctgaggtgacccaactcattaccataagaatttc	1269
Db	1092	CCCAAGATGTCGAATGCTGCTCCCAAACTCCCCACCTTTTGTACCATAGGAGATTC	1151
QY	1270	ctgaatgggtcactgaggtattcattcctgggttgatgctctcaccgtatgcacatg	1329
Db	1152	TTCCTACAAATCACCCTGGACTTCATCTGGCGATGGATGGCCCTCAAGAGGTTTACAATG	1211
QY	1330	aaatcattcattctgatg	1346
Db	1212	AGATAGTTACCATGATG	1228
RESULT	9		
LOCUS	AK010124		
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069D09, full insert sequence.	HTC	08-FEB-2001
ACCESSION	AK010124		
VERSION	AK010124.1	GI:12845346	
KEYWORDS	C&P trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (sites)		
TITLE	Carninci, P. and Hayashizaki, Y.		
JOURNAL	High-efficiency full-length cDNA cloning		
REFERENCE	Methods Enzymol. 303, 19-44 (1999)		
AUTHORS	2 (sites)		
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
REFERENCE	Genome Res. 10 (10), 1617-1630 (2000)		
AUTHORS	20499374		
TITLE	3 (sites)		
JOURNAL	Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.		
REFERENCE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		

Qy	670	atgtcgctattccagggaaccaccattggtcttattgcattttccaccatgccagagc	729
Db	549	ATGTGTGTGCTACTTCAGGCGACCACTACTCGGTTTTATTGCTTTTCTACCAATCTGCTC	608
Qy	730	tggctcgaataacaaatgtatttcttagcaccatagccacgtttaagcatgcaa	789
Db	609	TGGCTAAATAATCAAGAGGTTTTATGCAATAGCTCCAGTTGCTACTGTGAAGTATACAG	668
Qy	790	aaagcccccggaccacaaatttggctgcccagatgatgatcaagsgattgttgcca	849
Db	669	AAAGTCCCTTTAAAAAGATTTCATTATTCCTAAGTTCTCTCAAGGTGATATTGGTA	728
Qy	850	aaaaagaattctctgatacagaccagatctctcagacaaatttcttattacotttggcc	909
Db	729	ACAAAATGTTTCATGCCGCCACAACACTACATAGATCAATTTCTTGGTACGGAAGTGTCTAC	788
Qy	910	agtgattcttgaccagattgttagtaatacatgttacttctgggtgagattcaaccca	969
Db	789	GGGAGCTGCTAGATCTTCTCTGCAGCAACGGTTTATTTCATCTCTGTGGATTGACAAGA	848
Qy	970	acaatatgaacatgagccgagcaagtgtatatgctgccacactctgtggaacatctg	1029
Db	849	AAAATTTAAATGTGAGTCGCTTTGATGTGTATCTAGGCGATAATCCAGCAGAACATCTA	908
Qy	1030	tgcataattctacatgagccagcagtgaaattctgtgaaactccggcatttgact	1089
Db	909	CTCAAGACCTTTTCCATCTGGCACAGCTGTGTAATCTGGGAAGCTTCAAGCCTATAACT	968
Qy	1090	ggggagctgagaccacaaatctggaataatgcaatcagccaactccgttaaggtacagag	1149
Db	969	GGGAATGCTCATACAGAACATGTTACACATACATAAGAAACCCCTCTACTATGATG	1028
Qy	1150	tcagagatatgacggtccctacagcaatgtggacagaggtcagagactgggtttcacaatc	1209
Db	1029	TGTCAGCCATGACCGTGCCAAATGTCAGTGTGGAAAGGTGGCCATGACATCTCGCTGATC	1088
Qy	1210	cagaagacgtgaaatgctctctgaggtgacccaacctcatcaccataagaatttc	1269
Db	1089	CCCAAGATGTCGCAATGCTGCTCCCAAACTCCCAACCTTCTGTACCATAAAGCAGATT	1148
Qy	1270	ctgaatgggtcagctgagattctctgaggttggatgctctcaccgtatgacaatg	1329
Db	1149	TTCCCTACAATCAGCTGGAGTTCATCTGGCGATGGATGGCCCTCAAGAGGTTTACAATG	1208
Qy	1330	aaatcatccatctgagtg	1346
Db	1209	AGATAGTTACCATGATG	1225
RESULT 11			
AK009413			
LOCUS			
DEFINITION			
Mus musculus adult male tongue cDNA, RIKEN full-length enriched			
library, clone:2310020E22, full insert sequence.			
AK009413			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
CAP trapper.			
Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,			
clone_11b:RIKEN full-length enriched mouse cDNA library			
clone:2310020E22.			
Mus musculus			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
1 (sites)			
Carninci, P. and Hayashizaki, Y.			
High-efficiency full-length cDNA cloning			
Methods Enzymol. 303, 19-44 (1999)			
REFERENCE			
2 (sites)			
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
Normalization and subtraction of cap-trapper-selected cDNAs to			
prepare full-length cDNA libraries for rapid discovery of new genes			
Genome Res. 10 (10), 1617-1630 (2000)			
JOURNAL			

QY	970	acaatataaacatgacccgagcaagatgtaatgctgcaccacactcttcgtaacatctg	1029
Db	852	AAAACCTTAAATGCTGAGTGCCTTTGATGTGTTCTTAGGCCATTAATCACCAGGAACATCTCA	911
QY	1030	tgcacaatattctacactgtgagccagagagtgtaattctcgtgtaactccggagcatttgat	1089
Db	912	CTCAAGACCTTTTCCACTGGCGACAGCTTGCTGAATATCTGGGAAGCTTCAAGCCTTAACCT	971
QY	1090	gggggagagagbaccacaaaattctggaanaattgcattcagcccaactctctaagctacagag	1149
Db	972	GGGGAAAGTCCATTACAGAACTATTACACTCAACAAACAACAAAGCGCTCCCTATCTATGATG	1031
QY	1150	tcaagatatatgacgctccctacacgaatgtgacagagagtcgaagcagcttcataact	1209
Db	1032	TGTCAGGCATATGACCGTGGCAATTGCAAGTGTGGAAAGCGTGGCCATATACATCTCGCTGATC	1091.
QY	1210	cagaagaagctgtaaaaatgctgctctctctgaggtgacccaactatctacataagaatatlc	1269
Db	1092	CCCAAGATGTGTCAGATGCTGCTCCCAAACTCCCAACCTTCTGTATCATTAGAGGATGTC	1151
QY	1270	ctgaatggagctcagatgatttactatctgaggttgtagatgcctcactaccgtatgtcaatg	1329
Db	1152	TTCCCTACATCATCTGACTTTCATCTTGCGGAGTAGATGCGCTTAAGAGTTTACAATG	1211
QY	1330	aaatcatccatctgcatg	1346
Db	1212	AGATAGTTTACCATGATG	1228
RESULT	14		
LOCUS	AK009523		
DEFINITION	AK009523	1346 bp mRNA	HTC.
ACCESSION	AK009523	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310028109, full insert sequence.	08-FEB-2001
VERSION	AK009523.1	GI:12844369	
KEYWORDS	CAP trapper.		
SOURCE	Mus musculus (Strain: C57BL/6J) adult male tongue cDNA to mRNA, clone:110:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (sites)		
REFERENCE	Carninci, P. and Hayashizaki, Y.		
REFERENCE	High-efficiency full-length cDNA cloning		
REFERENCE	Methods Enzymol. 303, 19-44 (1999)		
REFERENCE	2 (sites)		
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
REFERENCE	Genome Res. 10 (10), 1617-1630 (2000)		
REFERENCE	3 (sites)		
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagase, S., Sasaki, H., Carninci, P., Kono, H., Akiyama, Y., Nishi, K., Kitsuma, T., Tashiro, H., Itoh, M., Kikuchi, N., Ishii, T., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaiguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanabe, T., Matsura, S., Okazaki, Y., Muramatsu, M., Inoue, Y. and Hayashizaki, Y.		
REFERENCE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
REFERENCE	Genome Res. 10 (11), 1757-1771 (2000)		
REFERENCE	4 (sites)		
REFERENCE	The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.		
REFERENCE	Functional annotation of a full-length mouse cDNA collection		
REFERENCE	Nature 409, 685-690 (2001)		
REFERENCE	5 (bases 1 to 1346)		
REFERENCE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,		

TITLE
JOURNAL

Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiroaka,T., Horii,F., Imotoani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Kono,H., Koude,M., Koya,S., Kuribara,C., Matsuyama,T., Miyazaki,A., Nishikawa,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Oshida,T., Ono,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,I., Tagawa,A., Takahashi,F., Tanaka,T., Yoshimura,T., Toyota,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.go.jp), URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAAGGAATCCCAAGACCTCTTTTCTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAAATTCCTCAGTTAATTAAATAATCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

Location/Qualifiers
L . 1346
/organism="Mus musculus"
/strain="C57BL/6j"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1895759"
/db_xref="MGI:1914967"
/clone="2310028109"
/sex="male"
/tissue-type="tongue"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53. 1240
/note="putative"
/codon_start=1
/protein_id="BAB26338.1"
/translation="MWLLVTSYLSAFGAHGLFGKLPKNPEANNMVSONITWGYPSSEEEVETDEGILIGVRYIRPYKKNSNEDKRPAVAYOHDLHSATSMTNPLNSLAFLADAGVDWLINSGNTSRRKNVYSPPSVAPMSDPBKAKYDLPIDRVQGTGOEKIRHVGHSGOCTTIFIASTRPALAKIKRYALAPAIVATYTESPKKISLIKYLKHVITGNMFMPHPNLLDQFLGTVCSRELLDLCSNALFTFCGFQRKLNVSRFVFLFNPGSTPDTLFMAWDLAGSKGLDAIYWGSPLOMLHYNQTPRYDVDSANTVTI AAVNGGHDIADPDVAMLLPKLPNLILYHEILIPYNHLDETWMADAPOEVYEIVTM AED"

BASE COUNT 395 a 289 c 283 g 379 t
ORIGIN

Query Match 22.4%; Score 458.6; DB 192; Length 1346;
Best Local Similarity 63.6%; Pred. NO. 2:4e-106;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

250 accagaagcattatgatatagtgaatcatcccaacatcaaggctaccctcgtaggg 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 132 ACCCTGAAGCAAACGTGATCTTAGTCAGATGATTAAGTACTGGGATATCAAGTGAGG 191
||| |

310 aaatgaagtcgaactaaagatgggtatcatcctttctgttaaacagagatctccgagacc 369
||| |

Db 192 AATATGAAGTCTTACTAGAGATGGCTACATCTCTGGGGTCTATAGAAATTCCTTATGGA 251
QY 370 tagtgcaacctaaagacaggttccaggcctgtgtgttactgtcagcatggcctagtgtg 429
Db 252 AGAAAAATCTGAGATAATCGCAAGACAGCTGTGGCATATTTTGTGAGCATGGTTGATTG 311
QY 430 gagggtcgtacgaactggattttccaaactgcgcaacaatagcctgggttcattctcggcag 489
Db 312 CATCAGCCCAAACTGGATTCAGAAATCGCCAAACAGAGCTGGCCCTTCATCTAGGACG 371
QY 490 atgtcgtgtttgacgtgtggaagggaacagcagggaacccctgggtctcgaacaacaca 549
Db 372 ATGCTGGCTATGATGTGTGGCTGGGAACAGCTCGAGGGAATACATGGTCCCGAAATG 431
QY 550 agacactctccatagacaagaatgagttcgtggcttcagtttatgatgatgagtggttaggt 609
Db 432 TATACATATTCACAGACTCAGTTGAATCTCTGGCTTTTTCAGCTTTTGTGAAATGGCTAAT 491
QY 610 ttgacctctcgtcagtgataaactttatttgcagaaacgggccaagaagaatctatt 669
Db 492 ATGACCTTCACGCCACCATAGACTTCATTTGACAGAAACTGGCAAGAGAGATACACT 551
QY 670 atgtcggctattcacaggccacacacatggcgtttattgcatcttccaccatgccagagc 729
Db 552 ATGTTGGTCACTCTCAGGCCACCATATCGGTTTATTGGCTTTCTACCAATCTGCTC 611
QY 730 ttgctcagaataatcaaatgtattttgttttagcaccatagccactgttaagcatgcaa 789
Db 612 TGGCTAAAAAATCAAGAGGTTTATGCAATAGCTCCAGTCTACTGTGAAGTATACAG 671
QY 790 aaagcccgccgacaaatatttctgtcgcagatatgatataagggattggttgca 849
Db 672 AAAGTCCCTTTAAAAAGATTTTCCATTTTCTTAAAGTTTCTTCAAGGTGATTTGGTGA 731
QY 850 aaaaagaattctgtatcagaccagacttctcagacaacttctgtatttaccctttgtggc 909
Db 732 ACATAATGTTTCATGCCCCACACTACTTAGATCAATTTCTGGTAGGAGTGTGCTCAG 791
QY 910 aggtgattcttgatcagattgttagtaatatcatgttactcttgggtgattcaacaca 969
Db 792 GGGAGCTGTAGATCTCTCTCAGACAGCGCTTTATTATCATCTCTGTGGATTGACAAGA 851
QY 970 acaatatgaactagccagcagcaagtgtatgtatgtcgcacacactcttctggaacatctg 1029
Db 852 AAAACTTAAATGTGATCGCTTTGATGTGTATCTAGGGCATATCCAGCAGGAACATCTA 911
QY 1030 tcaaaaattctcactgagccagccagtgaaattctggtgaactccgggcatttgact 1089
Db 912 CTCAGACCTTTTCCACTGGGCACAGCTTGTGTAATCTGGGAGCTTCAAGCCTATAACT 971
QY 1090 ggggagtgagacaaatactggaataatgcaatcagccaactcctgttaagggtacagag 1149
Db 972 GGGGAAGTCCATTACAGAACATGTTACACTACAATCAGAAAAACGCTCCTACTATGATG 1031
QY 1150 tcagagatatgaggtccctcagacaatgtggacaggaaggtcagagactggttcaaatc 1209
Db 1032 TGTCAACCTTACCGTGGCAATTCGAGTGTGGAACGGTGGCCATGACATCTCTGGTGTGATC 1091
QY 1210 cagaagacgtgaaatcgtcctcctcgtgagtgaccaaactcattaccataagaatttc 1269
Db 1092 CCCAAGATGCGAATGCTGCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC 1151
QY 1270 ctgaatgggctcagctggatttcattcattgggtttggtgctcctcaccgtatgtacaatg 1329
Db 1152 TTCCCTTACAATCACCTGGACTTCATCTGGCGATGGATCGGCTCAAGAGGTTTACAATG 1211
QY 1330 aaatccatctatgatg 1346
Db 1212 AGATAGTTACCATGATG 1228

RESULT 15
AK009729

LOCUS AK009729 1346 bp mRNA HTC 08-FEB-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310040L03, full insert sequence.
ACCESSION AK009729
VERSION AK009729.1 GI:12844702
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library
clone:2310040L03.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods Enzymol. 303, 19-44 (1999)
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE 20499374
AUTHORS 3 (sites)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Kikuchi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Farada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Ozawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Okazaki,Y., Muramatsu,M., Inoue,Y. and Hayashizaki,Y.
TITLE FIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE 20530913
AUTHORS 4 (sites)
The RIKEN Genome Exploration Research Group Phase II Team and FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5 (bases 1 to 1346)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
TITLE Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2001, 22:38:56 ; Search time 97.86 Seconds
(without alignments)
2454.893 Million cell updates/sec

Title: US-09-333-159-46
Perfect score: 1269
Sequence: 1 atgttggaacctgtgcag.....gacgggtgagcgccgtattg 1269

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appli
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appli
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appli
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appli
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appli
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appli
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appli
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appli
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appli
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appli
11	46.4	3.7	43676	3 US-09-356-952-12	Sequence 12, Appli
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appli
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appli
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 154, App
15	36.6	2.9	3282	1 US-08-276-852-159	Sequence 169, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
17	36.6	2.9	3282	1 US-08-899-575-159	Sequence 169, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
19	36.6	2.9	3282	1 US-08-899-575-159	Sequence 169, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
21	36.6	2.9	3282	5 PCT-US95-08743-159	Sequence 169, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
23	36.6	2.9	13254	1 US-08-276-852-170	Sequence 170, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
25	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
27	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App

28	36.6	2.9	13254	5 PCT-US95-08743-156	Sequence 156, App
c 29	36.6	2.9	13254	5 PCT-US95-08743-170	Sequence 170, App
30	33.6	2.6	8392	1 US-08-080-255-6	Sequence 6, Appli
31	33.6	2.6	8392	3 US-08-465-713-6	Sequence 6, Appli
32	33.6	2.6	8392	5 PCT-US93-05857-6	Sequence 6, Appli
c 33	32.6	2.6	7218	1 US-08-232-453-14	Sequence 14, Appli
c 34	32	2.5	2197	1 US-08-233-005-3	Sequence 3, Appli
c 35	32	2.5	2197	1 US-08-428-943-3	Sequence 3, Appli
c 36	32	2.5	2197	3 PCT-US95-04858-3	Sequence 3, Appli
c 37	32	2.5	2197	5 PCT-US95-04858-3	Sequence 3, Appli
38	32	2.5	8342	3 US-08-545-860D-63	Sequence 63, Appli
39	32	2.5	8342	5 PCT-US94-04496-63	Sequence 63, Appli
40	31	2.4	3212	3 US-08-673-814-4	Sequence 1, Appli
c 41	30.4	2.4	1529	1 US-08-336-778-1	Sequence 1, Appli
c 42	30.2	2.4	4982	3 US-08-699-103B-1	Sequence 1, Appli
c 43	30.2	2.4	10079	2 US-08-476-866-20	Sequence 20, Appli
c 44	30	2.4	3639	2 US-08-737-524B-26	Sequence 26, Appli
45	30	2.4	5178	2 US-08-474-169-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-227-108-2
Sequence 2, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 accagagcattcatgaatattagtgaaatcatcccaacatcaaggctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGATACCCAGCTGAG 85
QY 185 aatatgaagtcgaactgaagatgggtatatctctttctgttaacagagatccctgggcc 244
DB 86 AATATGAAGTGTGACCGAAGAGCGTTATATCTTGGGATCGACAGAAATTCCTTATGGGA 145
QY 245 tagtgaacacctaagaagacaggttccagcctgtgtgttactgcagcatggcctagt 304
DB 146 GGAAAAATTTCAGAGATATAGCGCGGAGACCTGTGCAATTTTGCAACAGGTTTGTCTG 205
QY 305 gagggtgtagcaactgatttccaaactgcccacaaatagcctgggcttctattctggcag 364
DB 206 CATCAGCCCAAACTGGATCTCCAACTGCCAACACAGCCTGGCTTCTATCTGTGCCG 265
QY 365 atgctggtttgacgtgtggtatggatgggggaaacagcgaggggaaacacaca 424
DB 266 ACGCCGGGTACGACGTGTGGCTGGGGAACAGCAGCGGCAACACTGGCGCAGGGAATC 325
QY 425 agacactctccatagacaaagatgagttctggctttcagttatgatgagatggctagt 484
DB 326 TGTACTACTGCCCGACTCCGTGGAATCTGGGCTTTTCAGCTTTGACGAGATGGCTAAAT 385
QY 485 ttgacctctcctgagtgataaaactttatttgcagaaacgggcccaggaagaaatctatt 544
DB 386 ATGACCTTCCCGCCACCATTTGACTTCATCTTGAGAAACGGGACAGACAGCTACT 445
QY 545 atgtcggtattacagagggcaccacacatgggctttatttgcatcttccaccatccagagc 604
DB 446 ACGTTGGCCATTCCACGGGACCACTTCTTGTATCAATTTCTCGCCACGAGGTGCTCC 505
QY 605 tggctcagaataaataatttcttcttagcaccatagcactgttaacatcaaa 664
DB 506 TGGGAAACGGATCAAACTTCTATGATAGTCTCCGTTGCCACCGTGAAGTACACCG 565
QY 665 aaagcccgaggaccataattttgtgtgcagatgatgatcaaggatgtttggca 724
DB 566 AAACCTGTTAAACAACTCATGCTCGCTTCTTCAAGCTTATATTGGAA 625
QY 725 aaaaagaattctctgatacagaccagattctcagacaactgtttatttacccttgggcc 784
DB 626 ACAAATATTCTACCCACACACTTCTTGTATCAATTTCTCGCCACGAGGTGCTCC 685
QY 785 aggtgattctgacagattgttagtaataatcatgatttacttctgggtggattcaacca 844
DB 686 GCGAGCGGTGGATCTCTTGCAGCAACGCCCTGTGTATCATTTGTGGATTTGACACTA 745
QY 845 acaatatgaacatgagcagagcaagtgatgctgtgccacactctgtggaacatctg 904
DB 746 TGAACCTTGAACATGAGTCGCTTGGATGTGTATCTGCATATCCAGCAGGAACATCG 805
QY 905 tgcataatattctacatgagcagcagcagtgaaattctgtggaactccggcatttgact 964
DB 806 TTCAAGATGCTGCTCAGCTGCTCCAGCTGTTAAAGTCTGGGAAGTTTCAAGCTTTGACT 865
QY 965 ggggagtgagacaaaactggaataatgcaatcagcaactccctgtaaggatagag 1024
DB 866 GGGGAACCCGATTCAGAAATGATGACATATCATAGAGCATGCTTCCCTTACTACAA 925
QY 1025 tcagagatgacggttccctacagcaatgtggacagaggttcagagactggctttcaaatc 1084
DB 926 TGACAGACATGATGTGCCAATCGCAGTGTGGAACGGTGGCAACGACTTGTGCGCCGACC 985
QY 1085 cagaagacgtgaaatgctgctctctgaggtgaccaactctatccataagaataatc 1144
DB 986 CTCAGATGTTGACCTTTTGTGCTTCCAAAGCTTCCCAATCTCATTTACCAACAGGAATC 1045
QY 1145 ctgaatgggctcagtgagttctctctgggtttggatgctctcaccgctatgtacaatg 1204
DB 1046 CTCCTTCAATCACTTGGATTTATCTGGGCCATGGATGGCCCCCTCAAGCGGTTTACAATG 1105

QY 1205 aatcatccatctgatg 1221
DB 1106 AATTTGTTTCCATGATG 1122
RESULT 2
US-09-073-674-2
; Sequence 2, Application US/09073674
; Patent No. 598189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2
Query Match 35.0%; Score 444.2; DB 2; Length 1137;
Best Local Similarity 62.8%; Pred. No. 5.6e-139;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
QY 125 accagagcattcatgaatattagtgaaatcatcccaacatcaaggctatccctgtgag 184
DB 26 ACCCTGAAGTGACCATGAATATAAGTCAGATGATCACTACTGGGATACCCAGCTGAG 85
QY 185 aatatgaagtcgaactgaagatgggtatatctctttctgttaacagagatccctgggcc 244
DB 86 AATATGAAGTGTGACCGAAGAGCGTTATATCTTGGGATCGACAGAAATTCCTTATGGGA 145
QY 245 tagtgaacacctaagaagacaggttccagcctgtgtgttactgcagcatggcctagt 304
DB 146 GGAAAAATTTCAGAGATATAGCGCGGAGACCTGTGCAATTTTGCAACAGGTTTGTCTG 205
QY 305 gagggtgtagcaactgatttccaaactgcccacaaatagcctgggcttctattctggcag 364
DB 206 CATCAGCCCAAACTGGATCTCCAACTGCCAACACAGCCTGGCTTCTATCTGTGCCG 265
QY 365 atgctggtttgacgtgtggtatggatgggggaaacagcgaggggaaacacaca 424

[illegible]

D	b	217	TATCTGTGCGCCTTAACCGAATTCTCTCATGGAGGAACAACCATTCCTGCAAAAGGTCGCCAA	276
Q	y	273	gccctgggttgtaactgcagcatgccctagttggaggTgctagcaactggatttccaacct	332
D	b	277	ACCAGTTGTCITCCCTGCACAACATGCTTGCTGGCAGATTCTTAGTAACCTGGGTACAAAACCT	336
Q	y	333	gcccaacaataactcgtgggttcctaactcgtggcagaatgctgggttttgacagtgtgatgggaaa	392
D	b	337	TGCCAACAGCACGCTGGGCTTCATTCITGCTGATGCTGTGGTTTTGGAGTGTGGATGGCAA	396
Q	y	393	cagcaggggaasgcgcctggtctcgaaaacacaagaacactccatagacaagaatgagtt	452
D	b	397	CAGCAGAGGAATACTGGTCTCGGAACATAGAACAACACTCTCAGTTCACAGATGAATT	456
Q	y	453	ctgggcttcagtttatgatgatggttaggtttgacacctccctgcagtgataaaactttat	512
D	b	457	CTGGGCTTTTCAGTTATGATGATGGCAAAATATGACCTACCAAGCTTCCAATTAACTTCAT	516
Q	y	513	tttgcagaaaaacggcccaggaagaaagatctattatgctcggtctattccagggcaccaaat	572
D	b	517	TCGTAATAAAACTGGCCAAAGAACTATATATGTGGTCTANTCTCAAGCCACCACTAT	576
Q	y	573	gggcttatttgcattttccaccatgcagagctggtccagaaaaatcaaaatgatttttgc	632
D	b	577	AGGTTTTATAGCATTTTCACAGATCCCTGAGCTGGCTGAAAAAGGATTAATAATGTTTTTTCG	636
Q	y	633	tttagaccaatagccactgttaagcatgcaaaaagccccgggaccaaatttttgttgtct	692
D	b	637	CTGGGCTCTGTGGCTTCGGTCGCCCTTGCTACTAGCCCTATGGCCAAATAGGACGAT	696
Q	y	693	gccagatatgatgacaaggatgttttggcaaaaagaattttctgTatcagaccagatt	752
D	b	697	ACCAGATCATCTCATTAAGGACTTATTTGGAGACAAAGAAATTTCTTCCCACAGAGTCGTT	756
Q	y	753	tctcagacaacttyttatttaacctttgtggccagagtgattcttgatcagattttgtagttaa	812
D	b	757	TTTGAAGTGGCTGGGTACCACCGTTTGGCACTCATGTCTATCTAAGGAGGCTCTGTGAAA	816
Q	y	813	tatcatgttactctctgggtggattcaacaccaacaatatgaacatgagccgagcaagtgt	872
D	b	817	TCCTGTGTTTCTCTGTGTGGATTTTANTAGAGAGAAATTTAATAATGCTAGAGTGGATGT	876
Q	y	873	atatgtgccacaactcttgtcggaaactctgtgcaaaatatctacactggagccaggc	932
D	b	877	ATATACAACAACATTTCTCTGCTGGAAC'TCTGTGCAAAACATGTTACACTGGAGCCAGGC	936
Q	y	933	agtgaattctgtgaaactccgggactttgactggggagtgagaccaaaaacttgaaaaa	992
D	b	937	TGTTAAATCCCAANAAGTTTCAAGACCTTTGACTTGGGGAAGCAGTGCCCAAGAAATATTTTCA	996
Q	y	993	atgcaatcagccaactctgtgaagTacagagtcagagatatgscggtccctacagcaat	1052
D	b	997	TTACAACCAAGAGTTATCCTCCCAACATACAATGTGAAGGACATGCTGTGCGGACTGCAGT	1056
Q	y	1053	gtggacagggggtcaggaactgctttc aaaatccagaagaactgtaaaatgctgctctctga	1112
D	b	1057	CTGGAGCGGGGTCAGAGCTGGCTTGCAGATGCTCTACGACGCTAATAATCTTACTGACTCA	1116
Q	y	1113	ggtgaccaactcatctacataagaatatctgaaatgggctcacgtggatttcactg	1172
D	b	1117	GATCACCACACTTGGTGTTCATGAGAGCATTCGGAAATGGGAGCATCTTGACTTCAATTG	1176
Q	y	1173	gggtttggatgctcctcacogtatgtacatagaatacatccatcctatgctgagcaggagga	1232
D	b	1177	GGGCTTGGATGCCCTTGGAGGCTTTATAATAAAATTTAATCTAATGAGGAATATCA	1236
Q	y	1233	gaccaaacct	1241
D	b	1237	GTGAAGACT	1245

RESULT
A26689
5

Mon Aug